

# **Phylogenetische Untersuchungen von BAHD-Acyltransferasen in Lamiaceen und Boraginacen**

Dissertation

zur

Erlangung des Doktorgrades

der Naturwissenschaften

(Dr. rer. nat.)

dem Fachbereich Pharmazie

der Philipps-Universität Marburg

vorgelegt von

**Johannes Benner**

**aus Herbornseelbach**

Herbornseelbach im Dezember 2014



# **Phylogenetische Untersuchungen von BAHD-Acyltransferasen in Lamiaceen und Boraginaccen**

Dissertation

zur

Erlangung des Doktorgrades

der Naturwissenschaften

(Dr. rer. nat.)

dem Fachbereich Pharmazie

der Philipps-Universität Marburg

vorgelegt von

**Johannes Benner**

**aus Herbornseelbach**

Herbornseelbach im Dezember 2014

---

*Vom Fachbereich Pharmazie*

*der Philipps-Universität Marburg als Dissertation am \_\_\_\_\_ 2014 angenommen.*

*Erstgutachterin: Prof. Dr. M. Petersen*

*Zweitgutachter: Prof. Dr. M. Keusgen*

*Tag der mündlichen Prüfung: 17. Dezember 2014*



*Wer viel fragt, kriegt viel Antworten!*

# Inhaltsverzeichnis

|   |           |
|---|-----------|
| <b>1. Einleitung.....</b>   | <b>1</b>  |
| 1.1. Phylogenie .....   | 1         |
| 1.2. Genduplikation .....   | 4         |
| 1.3. Sekundärmetabolite .....   | 6         |
| 1.4. BAHD-Acyltransferasen .....  | 9         |
| 1.5. Die Hydroxycinnamoyltransferasen RAS, HST und HQT .....  | 10        |
| 1.6. Rosmarinsäure .....  | 12        |
| 1.7. Biosynthese der Rosmarinsäure .....  | 13        |
| 1.8. Vorkommen der Rosmarinsäure im Pflanzenreich .....   | 16        |
| 1.9. Phylogenie der Lamiaceae .....   | 18        |
| 1.10. Phylogenie der Boraginaceae .....   | 22        |
| 1.11. Zielsetzung .....   | 23        |
| <b>2. Material .....</b>  | <b>25</b> |
| 2.1. Geräte .....   | 25        |
| 2.2. Chemische Grundsubstanzen .....  | 26        |
| 2.3. Vektorkarten .....   | 28        |
| 2.4. Primer .....   | 30        |
| 2.5. Molekularbiologische Reagenzien und Kits .....   | 31        |
| 2.6. Bakterienstämme .....  | 31        |
| <b>3. Methoden .....</b>  | <b>32</b> |
| 3.1. Klonierung von RAS- und HST-PCR-Amplifikaten aus Pflanzenmaterial .....                          | 32        |
| 3.1.1. Isolierung von genomischer DNA aus Pflanzenmaterial .....                                      | 32        |
| 3.1.2. Isolierung von RNA aus Pflanzenmaterial .....  | 33        |
| 3.1.3. cDNA-Synthese .....  | 34        |
| 3.1.4. Polymerase Kettenreaktion (PCR) zur Amplifizierung von cDNA .....                              | 35        |
| 3.1.5. Isolierung von 5´- und 3´-Enden der kodierenden Gensequenzen eines Organismus – RACE-PCR ..... | 37        |
| 3.1.6. Agarose-Gelelektrophorese .....  | 40        |
| 3.1.7. DNA-Extraktion aus Agarosegelen .....  | 42        |
| 3.1.8. Ligation .....   | 43        |
| 3.1.9. Herstellung kompetenter Zellen .....   | 43        |
| 3.1.10. Transformation .....  | 44        |
| 3.1.11. Aufreinigung von Plasmid-DNA .....  | 46        |
| 3.1.12. Verdau mit Restriktionsenzymen .....  | 47        |
| 3.1.13. Photometrische DNA-Quantifizierung .....  | 48        |
| 3.1.14. DNA-Sequenzierung .....   | 49        |
| 3.1.15. Kryokonservierung von Prokaryoten .....   | 49        |
| 3.2. Expressionsversuche der gefundenen RAS und HST cDNAs .....                                       | 50        |
| 3.2.1. Expression der cDNA .....  | 50        |
| 3.2.2. Optimierung der Expression .....   | 51        |
| 3.2.3. Herstellung von Proteinrohextrakt .....  | 53        |
| 3.2.4. Synthese der CoA-aktivierten Kaffeesäure (Caf-CoA) bzw. 4-Cumarsäure (pC-CoA) .....            | 54        |
| 3.2.5. Proteinreinigung über His-Tag Affinitätschromatographie .....                                  | 55        |
| 3.2.6. Umpuffern der Elutatfraktionen über Größenausschlusschromatographie (SEC) .....                | 56        |
| 3.2.7. Umpuffern und Konzentrierung durch Dialyse .....   | 57        |
| 3.2.8. Proteingehaltsbestimmung nach Bradford .....   | 58        |
| 3.2.9. SDS-PAGE .....   | 59        |

|         |  |     |
|---------|--|-----|
| 3.2.10. | <i>TCA-Fällung zur Aufkonzentrierung</i>   | 62  |
| 3.3.    | <i>Bestimmung der Enzymaktivität von RAS, HQT und HST</i>  | 62  |
| 3.3.1.  | <i>Enzymtests mit pHPL, Chinasäure und Shikimisäure</i>  | 65  |
| 3.3.2.  | <i>Extraktion der Produkte aus Enzymtests</i>  | 67  |
| 3.3.3.  | <i>Hochleistungsflüssigkeitschromatographie (HPLC)-Analyse</i>   | 67  |
| 3.4.    | <i>Statistische und phylogenetische Methoden</i>   | 69  |
| 3.4.1.  | <i>Chi<sup>2</sup>-Test</i>  | 69  |
| 3.4.2.  | <i>Berechnung der Intronspleißstellen mit NetPlant Gene Server</i>                                       | 69  |
| 3.4.3.  | <i>BAHD-Acyltransferasen in der pubmed-Datenbank</i>   | 70  |
| 3.4.4.  | <i>Phylogenetische Analyse von BAHD-Acyltransferasen mit MEGA 5.0</i>                                    | 70  |
| 4.      | <i>Ergebnisse</i>  | 73  |
| 4.1.    | <i>Isolierung von RAS-ähnlichen Teilsequenzen aus verschiedenen Pflanzenarten</i>                        | 73  |
| 4.1.1.  | <i>Alignment der Aminosäuresequenzen der amplifizierten RAS-Teilstücke</i>                               | 77  |
| 4.1.2.  | <i>Alignment der DNA-Sequenzen der amplifizierten RAS-Teilstücke</i>                                     | 82  |
| 4.2.    | <i>Isolierung von HST-ähnlichen Teilsequenzen aus verschiedenen Pflanzenarten</i>                        | 88  |
| 4.2.1.  | <i>Alignment der Aminosäuresequenzen der amplifizierten HST-Teilstücke</i>                               | 90  |
| 4.2.1.  | <i>Alignment der DNA-Sequenzen der amplifizierten HST-Teilstücke</i>                                     | 96  |
| 4.3.    | <i>Isolierung einer HST-ähnlichen Gesamtsequenz aus Symphytum officinale</i>                             | 104 |
| 5.      | <i>Diskussion</i>  | 115 |
| 5.1.    | <i>Methodik</i>  | 115 |
| 5.2.    | <i>Gewinnung von putativen RAS-Teilsequenzen aus verschiedenen Pflanzenarten</i>                         | 116 |
| 5.3.    | <i>Gewinnung von putativen HST-Teilsequenzen aus verschiedenen Pflanzenarten</i>                         | 121 |
| 5.4.    | <i>Gewinnung einer putativen HST-Volllängensequenz aus Beinwell (Symphytum officinale)</i>               | 132 |
| 5.5.    | <i>Erstellung eines phylogenetischen Baums mit den gewonnenen Teilsequenzen der verwendeten Pflanzen</i> | 135 |
| 6.      | <i>Zusammenfassung</i>   | 144 |
| 7.      | <i>Ausblick</i>  | 146 |
| 8.      | <i>Literaturverzeichnis</i>  | 147 |
| 9.      | <i>Abkürzungen</i>   | 157 |

# 1. Einleitung

## 1.1. *Phylogenie*

Der Grundstein für die Ökologie und auch Evolutionsforschung wird durch die Ergebnisse der Systematik gelegt. Die phylogenetische Systematik (Phylogenetik) ist ein Überbegriff für die Analyse von Verwandtschaftsbeziehungen auf molekularer Ebene, um eine überprüfbare Einordnung von Organismen in einen Stammbaum zu erzielen (Weberling und Schwantes, 2000). Die Phylogenese (Phylogenie) schafft die Möglichkeit, auch mit faktenbasierten Berechnungen die historische Entstehung von Arten und Artgruppen innerhalb von Stammbäumen abzubilden (Wiesmüller et al., 2002).

Für die Analyse von molekularen Merkmalen bei Pflanzen können die Genome des Kerns, der Chloroplasten und der Mitochondrien verwendet werden (Pichersky und Gang, 2000). Dabei tragen phylogenetische Analysen von Genen, die in bestimmten Biosynthesewegen beteiligt sind, in immer wichtigerem Maße zur Erforschung von neuen Substanzen und neuen Genen bei. Dies wird ermöglicht durch komplexe phylogenetische Werkzeuge (siehe 2. Methoden), die nur in Kombination ein schlüssiges Ergebnis präsentieren (Holder und Lewis, 2003). Weiter ist die phylogenetische Analyse ein Werkzeug, das durch die Kombination mit anderen Ergebnissen neue Einsichten eröffnet. Die Verzahnung von phylogenetischen Ergebnissen mit z.B. sekundären Pflanzeninhaltsstoffen oder morphologischen Merkmalen macht diese erst besonders aufschlussreich (Baltisberger, 2009).

Ein Weg, die Evolution einer charakteristischen Eigenschaft eines Organismus zu verstehen, ist eine taxonomische Ordnung der Eigenschaft über einen phylogenetischen Stammbaum zu erzeugen. Der phylogenetische Stammbaum zeigt eine evolutionäre Verwandtschaft von Organismen (z.B. Evolution der Lamiaceae anhand des *rbcl* Gens (siehe 1.5 Phylogenie der Lamiaceae)), welche eine charakteristische Eigenschaft jeweils besitzen oder teils auch nicht besitzen. Diese neuen Ergebnisse und sich daraus ableitende Hypothesen stehen jedoch nicht selten im Widerspruch zu Hypothesen aufgrund morphologischer Zuordnungen (Baltisberger, 2009).

Molekulare Merkmale besitzen gegenüber morphologischen Merkmalen wesentliche Vorteile. Die Erkennung molekularer Merkmale ist im Gegensatz zur Erfassung morphologischer Merkmale objektiver. Voneinander unabhängige DNA-Analysen

kommen grundsätzlich zum gleichen Ergebnis, während die Eingruppierung nach morphologischen Merkmalen subjektiv fehlerbehaftet sein kann (Schaal und Hayworth, 1998). Aus dem Vergleich homologer DNA-Sequenzen lässt sich die verwandtschaftliche Entfernung bestimmen, da die Mechanismen der Vererbung und der Merkmalsmutationen bekannt sind. Dabei ist jedoch zu beachten, dass stark abweichende Mutationshäufigkeiten und Mutationsgeschwindigkeiten in verschiedenen Genen/Organismen auftreten, die zu Fehlinterpretationen der verwandtschaftlichen Verhältnisse führen können. Der Vergleich homologer DNA-Sequenzen kann auch zu Diskrepanzen führen, die im Blick auf morphologische Merkmale entstehen können. Der Vergleich homologer DNA-Sequenzen eröffnet jedoch das größte Verständnis von evolutionären Zusammenhängen und Prozessen (Soltis und Soltis, 2003).

Phylogenetische Analysen beginnen mit der Bestimmung von Homologien. Dabei wird eine Auswahl an Nukleotidsequenzen oder Aminosäuresequenzen miteinander verglichen und die Gene für die spätere phylogenetische Analyse ausgewählt. Dieses *alignment* und die folgende Auswahl sind maßgeblich für die spätere Berechnung des phylogenetischen Stammbaums (Nei und Kumar, 2000).

Gene sind homolog, wenn sie von derselben Urart abstammen. Dabei bedeutet „homolog“ nicht, dass sie „gleich“ sind. Ist ein Protein B zu 90% mit Protein C homolog, bedeutet dies, dass 90% der Nukleotidsequenz beider Gene mit der Nukleotidsequenz eines Urgens A identisch sind. Die restlichen 10% sind evolutionär voneinander unabhängig (Reeck et al., 1987). Homologe Gene können paralog oder ortholog voneinander abhängen. Paraloge Gene sind durch Genduplikation entstanden (siehe 1.2 Genduplikation). Orthologe Gene stammen von einem Urgen des letzten gemeinsamen Verwandten (*ancestor*) der zu vergleichenden Spezies ab und sind durch verschiedene Spezialisierungen dieses Urgens entstanden (Koonin, 2005).

Die im Weiteren durchgeführten phylogenetischen Analysen sind statistische Auswertungen der ausgewählten Daten. Die eigentliche Gensequenz, die vor Beginn der evolutionären Veränderungen vorlag, ist nicht bekannt. Der phylogenetische Stammbaum muss daher berechnet werden. Den Berechnungen liegt die Beobachtung zu Grunde, dass die Ähnlichkeit der zu vergleichenden Organismen Ausdruck ihrer Verwandtschaft ist (Doolittle, 1999).

Es wird ständig an verbesserten Modellrechnungen geforscht, die mit größter Wahrscheinlichkeit die evolutionären Zusammenhänge der Daten verknüpfen. Dafür wurden eine Reihe von Berechnungsmodellen für die Generierung eines phylogenetischen Stammbaums entwickelt (zum Beispiel Distanzmethode, Maximale Parsimonie, Maximale Wahrscheinlichkeit) (siehe 2. Methoden).

Die verfügbaren Methoden sind alle als vereinfachte Modellrechnungen zu verstehen. Durch die Vereinfachung der wirklichen, hochkomplexen Evolutionsprozesse sollten mehrere Modelle untereinander auf ihre Aussagekraft hin verglichen werden. Die interpretierte evolutionäre Geschichte der untersuchten Daten sollte immer auf Robustheit und Signifikanz überprüft werden (Sneath und Socal, 1973). Auch hierfür sind Standardtestverfahren etabliert wie Goodness-of-fit-Indices, Decay-Index, Bootstrap-Index oder Skewness-Index.

Wird die Kartierung einer charakteristischen Eigenschaft um eine oder mehrere Eigenschaften ergänzt, können Korrelationen dieser Eigenschaften berechnet werden. Diese Analyse erlaubt eine Aussage, ob eine bestimmte Eigenschaft in Abhängigkeit einer zweiten Eigenschaft evolutionär entstanden ist. Anhand dieser Analysen kann zum Beispiel bestimmt werden, ob ein Metabolit in Abhängigkeit einer zweiten Substanz entsteht oder ob dieser Metabolit nur in Abhängigkeit bestimmter ökologischer Bedingungen zu finden ist (Maddison, 1990). Diese Ergebnisse sind wichtig für die Bewertung der Funktion von Sekundärmetaboliten und für das Verständnis über die Evolution von biologischen Synthesewegen. Über diese Methode konnte beispielsweise die Co-Evolution von verstärkter Produktion phenolischer Substanzen und verminderter Produktion von Cardenoliden in *Asclepias*-Arten nachgewiesen werden. Die Produktion von Cardenoliden und Flavonoiden sind hierbei über den Bedarf an Acetyl-CoA verbunden (Agarwal et al., 2008).

DNA-Analyse ist in den letzten 10 bis 20 Jahren eine zunehmend wichtige Methode, um die Phylogenie von Pflanzen neu zu ordnen oder zu ergänzen. Dabei ist die Kombination der DNA-Analyse mit der Auswertung über spezielle Computerprogramme (z.B. MEGA-Software) die beste Möglichkeit, um eine möglichst genaue taxonomische Ordnung zu erzielen (Kumar et al., 1994).

## **1.2. Genduplikation**

Die Evolution von Genen zur Bildung neuer Pflanzeninhaltsstoffe im Zuge der evolutionären Anpassung ist ein fortwährender Prozess. Nach heutiger Vorstellung ist die Genduplikation mit anschließender Divergenz die treibende Kraft für die Erschaffung neuer Sekundärstoffwechselgene (Pichersky und Gang, 2000). Dabei werden verschiedene Modelle postuliert, die jedoch noch nicht vollständig aufgeklärt oder bewiesen sind. Im Folgenden werden zwei anerkannte Modelle vorgestellt.

Eine sehr verbreitete Vorstellung, wie Genduplikation und Diversifikation funktionieren, ist, dass nach einer Duplikation der Organismus ein Gen besitzt, das die ursprüngliche Funktion übernimmt (Ohno, 1973). Das Tochtergen, auf welches kein natürlicher Selektionsdruck ausgeübt wird, bleibt redundant vorhanden. Da die redundanten Tochtergene keine essentiellen Funktionen übernehmen, akkumulieren Mutationen, und viele von ihnen werden nach Pseudogenisierung wieder aus dem Genom eliminiert, während andere Gene mit veränderter Enzymaktivität im Genom fixiert werden (Ohno, 1973). Ein neueres Modell nimmt den Gedanken der Genduplikation zwar auch auf, unterscheidet sich aber im Startpunkt: Genduplikation führt zu der Produktion von zwei Genen A' und A'', welche funktional unterschiedliche Proteine erzeugen. Jedoch wurden zuvor die beiden Funktionen von einem generalisierten Gen A kodiert. Nach der Genduplikation wird es möglich, dass sich die beiden aktiven Tochtergene auf jeweils einen Prozess spezialisieren. Im Verlauf der Evolution wird von den Tochtergenen nur noch oder größtenteils die spezialisierte Funktion ausgeübt (Hughes, 1994). Dieser theoretische Ansatz könnte die häufigen Überschneidungen von Genfunktionen erklären.

In der Literatur sind einige Beispiele beschrieben, die darauf hinweisen, dass Gene des pflanzlichen Sekundärstoffwechsels von Vorfahren aus dem Primärstoffwechsel abstammen (Bourgaud et al., 2001)

Ein Beispiel ist die Synthese von Pyrrolizidinalkaloiden. Diese sind Esteralkaloide, die aus der Necinbase und einem veresterten Säureanteil, den Necinsäuren, aufgebaut sind. Bedingt durch die Art und Anzahl der Necinsäuren und Veresterungen können Monoester, offenkettige Diester, makrozyklische Diester und makrozyklische Triester gebildet werden (Hartmann und Witte, 1995).

Die Necinsäuren der Pyrrolizidinalkaloide werden größtenteils aus den Aminosäuren Valin und Isoleucin abgeleitet. In wenigen Fällen dient auch Leucin als Ausgangsstoff

für die Synthese der Necinsäuren (Hartmann, 1999). Für die Synthese der Necinbase konnte der Schnittpunkt vom Primär- zum Sekundärstoffwechsel exakt identifiziert werden. Der Ausgangspunkt im Primärstoffwechsel ist das Arginin, welches über Putrescin zum Homospermidin umgesetzt wird (Ober und Hartmann, 1999). Homospermidin hat keine weitere Funktion im Primärstoffwechsel und ist die erste Substanz, die rein dem Sekundärstoffwechsel zugeordnet werden kann. Ermöglicht wird dieser Schnittpunkt zwischen Primär- und Sekundärstoffwechsel durch die Homospermidinsynthase (HSS, E.C. 2.5.1.45). HSS überträgt in einer NAD<sup>+</sup>-abhängigen Reaktion eine Aminobutylgruppe von Spermidin auf Putrescin. Das entstehende Homospermidin wird über weitere Reaktionsschritte zur Necinbase umgewandelt (Böttcher et al., 1993).

Die HSS hat ihren evolutionären Ursprung wiederum in der Desoxyhypusinsynthase (DHPS) des Primärstoffwechsels. DHPS katalysiert die posttranslationale Aktivierung des eukaryotischen Initiationsfaktors 5A (eIF5A) (Ober et al., 2003). Obwohl die Reaktionsprodukte von HSS und DHPS sehr unterschiedlich sind, katalysieren beide einen gleichen Reaktionsmechanismus: Sie transferieren durch eine NAD<sup>+</sup>-abhängige Reaktion die Aminobutylgruppe von Spermidin auf ein zweites Substrat. Phylogenetische Untersuchungen von 23 cDNA-Sequenzen von HSS und DHPS von verschiedenen bedecktsamigen Pflanzen zeigen, dass sich mindestens vier unabhängige HSS aus DHPS entwickelt haben (Reimann et al., 2004).

Trotzdem können Gene, die schon eine Funktion im Sekundärstoffwechsel übernommen haben, ihrerseits die Basis für die Gewinnung neuer Funktionen sein. In vielen Fällen gehören diese Gene zu großen Genfamilien wie den N- und O-Methyltransferasen, den Acyltransferasen oder den Cytochrom-P450-abhängigen Monooxygenasen. Gene aus dem Sekundärmetabolismus müssen keine Kopie bereitstellen, die eine essentielle Funktion des primären Metabolismus aufrecht erhält (Doolittle, 1995). Sie können auch ohne Duplikation mutieren (Pichersky und Gang, 2000).

Mitglieder großer Proteinfamilien, die wahrscheinlich durch Genduplikation und Diversifikation entstanden sind, weisen trotz deutlicher Sequenzunterschiede einige Gemeinsamkeiten auf, zum Beispiel besitzen BAHD-Acyltransferasen eine minimale Sequenzidentität von 10 bis 30% auf Aminosäureebene, jedoch kann die Existenz der stark konservierten Motive HxxxDG und DFGWG beobachtet werden (Yu et al., 2009). Auch die einheitliche dreidimensionale Struktur von Vertretern einer



Proteinfamilie ist ein Indiz für einen gemeinsamen evolutionären Ursprung (siehe Kapitel 1.5.1 BAHD-Acyltransferasen). Ähnliche Tertiärstrukturen können auch auf eine strukturelle Notwendigkeit für die biologische Funktion der BAHD-Acyltransferasen hinweisen (siehe Kapitel 5. Diskussion).

### **1.3. Sekundärmetabolite**

Pflanzen produzieren eine Fülle von Metaboliten mit geringer molekularer Masse. Mehr als 200000 Strukturen sind mittlerweile bekannt (Bresinsky et al., 2008). Nur wenige gehören zum primären Metabolismus, den alle Organismen teilen, der Rest wird dem Sekundärstoffwechsel zugerechnet. Per Definition ist damit ein Stoff, der nur durch eine oder wenige spezialisierte Pflanzenarten produziert wird, gemeint (Pichersky und Gang, 2000).

Die Biogenese der Sekundärmetabolite ist von der begrenzten Zahl der Primärmetabolite abzuleiten. Die Sekundärmetabolite können dabei aufgrund der Bausteine ihrer Biosynthese in drei übergeordnete Gruppen eingeteilt werden: stickstoffhaltigen Sekundärprodukte (z.B. Alkaloide, Amine), die Terpenoide (z.B. Carotinoide, Saponine) und die phenolischen Verbindungen (z.B. Flavonoide, Chinone) (Bourgaud et al., 2001), wobei auch Sekundärmetabolite existieren, die Komponenten aus mehr als einer dieser Gruppen enthalten.

Der Hauptteil der phenolischen Sekundärstoffe leitet sich aus dem Phenylpropanstoffwechsel ab. Die dafür notwendigen aromatischen Aminosäuren werden über den Shikimatweg bereitgestellt. Terpene hingegen werden über den Acetat-Mevalonat- und 3-Phosphoglycerat-Pyruvat-Weg synthetisiert (Bresinsky et al., 2008)

Pflanzen sind sesshafte Organismen, die vor potentiellen Gefahren nicht fliehen können. Sie besitzen auch kein „eigentliches“ Immunsystem (*innate immunity*), das sie vor Mikroorganismen schützen könnte. Somit mussten sie andere Strategien zum Überleben und zur Verteidigung entwickeln. Eine anerkannte Theorie ist, dass die Produktion von Sekundärmetaboliten als chemische Abwehr wirkt (Harborne et al., 1993).

Sekundärmetabolite sind, wie erwähnt, zahlreich und in allen höheren Pflanzen zu finden. Sie können nur als chemische Abwehr wirken, wenn sie in angreifenden Organismen molekulare Ziele negativ beeinflussen. Molekulare Ziele können dabei z.B. DNA, Proteine, Membranstrukturen oder Stoffwechselreaktionen sein (Wink und Schimmer, 1999). Es konnten bisher eine Vielzahl an spezifischen Interaktionen von Sekundärmetaboliten mit Proteinen (z.B. Strukturproteinen, Ionenkanälen, Rezeptoren oder Enzymen) nachgewiesen werden (z.B. Pyrethrine, Colchicin, Strychnin, Rotenon, Digitoxin etc.) (Bresinsky et al., 2008). Die Strukturen solcher Sekundärmetabolite scheinen im evolutionären Verlauf dahingehend geformt worden zu sein, dass sie möglichst perfekt die Struktur von endogenen Substraten wie z.B. Hormonen oder Neurotransmittern nachbilden (Harborne et al., 1993). Weit verbreitet sind aber auch Sekundärmetabolite, die mit mehreren molekularen Zielen unspezifisch reagieren. Zum Beispiel besitzen Tannine und andere Phenole durch ihre große Anzahl an phenolischen Hydroxylgruppen potentiell die Möglichkeit, Wasserstoffbrückenbindungen und ionische Bindungen mit allen Proteinen auszubilden. Die Proteine ändern dabei ihre Konformation, verlieren dadurch ihre Aktivität und können ihre Funktionen nicht mehr ausüben (Bresinsky et al., 2008).

Sekundärmetabolite sind als taxonomische Größe verwendbar und können einen starken phylogenetischen Bezug aufweisen. Die Anwesenheit oder Abwesenheit von Sekundärstoffen erlaubt in manchen Fällen eine phylogenetische Neuordnung von Pflanzen. Manche Sekundärstoffe werden von bestimmten Pflanzenarten nicht mehr gebildet, obwohl sie in Pflanzen, von denen sie abstammen, vorkommen. Das ungleiche Profil von Sekundärmetaboliten in einer Reihe von untereinander verwandten Pflanzenarten zeigt, dass die systematische Einordnung der chemischen Eigenschaften neue phylogenetische Analysen ermöglicht. Diese Einordnung ist ähnlich der traditionellen Charakterisierung über morphologische Eigenschaften nur durch das Auffinden, Charakterisieren und Interpretieren von Sekundärmetaboliten möglich (Nei und Kumar, 2000).

Die Synthese von spezifischen Sekundärmetaboliten wird als zusätzliche charakteristische Eigenschaft von Pflanzen betrachtet. Diese Eigenschaften wurden durch Selektion während der Evolution geformt und können somit für eine taxonomische Eingliederung Verwendung finden. Zum Beispiel kann die Synthese von Rosmarinsäure, welche im gesamten Pflanzenreich weit verstreut zu finden ist, für die Berechnung der evolutionären Zusammenhänge innerhalb der Familie der

Lamiaceae aufschlussreich sein (Petersen et al., 2009). Es ist jedoch sehr selten, dass ein einzelner Sekundärmetabolit zu einer durchgehenden phylogenetischen Gruppe zugeordnet werden kann. Innerhalb der Familie der Lamiaceae wird Rosmarinsäure hauptsächlich in der monophyletischen Unterfamilie der Nepetoideae gefunden. Einzelne Mitglieder aus der Unterfamilie der Teucroideae besitzen jedoch auch die Eigenschaft der Rosmarinsäuresynthese (Petersen et al., 2009).

Die Stoffgruppe der Iridoidglykoside stellt dagegen einen guten Marker für die Unterfamilie der Lamioideae dar. Aber auch hier gibt es Ausnahmen: *Nepeta cataria* und *Satureja vulgaris*, Mitglieder der Nepetoideae, enthalten viel Iridoide. In der zu Lamiaceae nahe verwandten Familie Verbenaceae sind Iridoide auch sehr stark vertreten. Dies ist wiederum ein Hinweis, dass es eine mit beiden Familien verwandte Urart gegeben haben könnte, in der die Synthese von Iridoiden gegeben war. Im Lauf der Evolution könnten dann die korrespondierenden Gene zur Synthese von Iridoiden bei Nepetoideae inaktiviert worden sein (Wink, 2003).

Andererseits gibt es strukturell hoch komplexe Moleküle, die gleichzeitig in nicht verwandten Pflanzenfamilien zu finden sind. Diese Moleküle können durch die breite Streuung nicht für eine evolutionäre Analyse verwendet werden. Zum Beispiel wurde Camptothecin, ein Topoisomerasehemmstoff, der zur Tumorbekämpfung eingesetzt wird, in den unterschiedlichsten, nicht nahe verwandten Pflanzenfamilien (Icacinaceae, Nyssaceae, Rubiaceae, Apocynaceae und Loganiaceae) gefunden (Lorence und Nessler, 2004)).

Ein großer Teil der Gene vieler Organismen kodiert Enzyme, die eine Funktion im Sekundärmetabolismus einnehmen. Zum Beispiel sind 9,9% des Genoms des marinen Bakteriums *Salinispora tropica* für die Produktion von Sekundärmetaboliten belegt (Udwary et al., 2007).

Diese Vielzahl und Diversität an Enzymen zur Produktion von Sekundärmetaboliten zeigt die Notwendigkeit von systematischen Ansätzen zur Entdeckung einzelner Gene mit dem Ziel der Bestimmung von ganzen Biosynthesewegen. Am Ende dieses Weges steht eine einzelne, teils hoch komplexe Substanz, deren funktionelle Eigenschaften sehr großen Einfluss auf die evolutionäre Entwicklung des entsprechenden Organismus hat (Bourgaud et al., 2001).

#### 1.4. BAHD-Acyltransferasen

BAHD-Acyltransferasen sind monomere Enzyme mit einem Molekulargewicht von 48-55 kDa. Die Abkürzung BAHD leitet sich aus den ersten biochemisch charakterisierten Enzymen der Familie ab: Benzylalkohol O-Acetyltransferase (BEAT), Anthocyan O-Hydroxycinnamoyltransferase (AHCT), Anthranilat *N*-Hydroxycinnamoyl/Benzoyltransferase (HCBT) und Desacetylvindolin 4-O-Acetyltransferase (DAT) (St. Pierre und De Luca, 2000). Die Funktion der BAHD-Acyltransferasen besteht darin, Acylreste von CoA-aktivierten Säuren auf Hydroxyl- oder Aminogruppen der Akzeptormoleküle zu übertragen (D'Auria, 2006). Sie sind aus ungefähr 445 Aminosäuren aufgebaut. Fast alle bekannten Mitglieder der BAHD-Familie enthalten mit dem an der Katalyse beteiligten HXXXDG-Motiv im mittleren Bereich des Proteins und dem in der Nähe des C-Terminus lokalisierten DFGWG-Motiv zwei hoch konservierte Bereiche (D'Auria, 2006). Ein drittes Motiv ist weniger stark konserviert und besteht aus den 13 Aminosäuren LSXTLXXXYYXXXG (Aharoni et al., 2000). Außerdem ist in vielen BAHD-Acyltransferasen ein Q-Intron (hinter der hoch konservierten Aminosäure Q = Glutamin) zu finden. Es sitzt 17 Aminosäuren vor dem konservierten HXXXDG-Motiv (St. Pierre und De Luca, 2000).

Anhand ihrer Aminosäuresequenz und ihrer Substratspezifität können die BAHD-Acyltransferasen in 5 Klassen eingeteilt werden (D'Auria, 2006). Alle in dieser Arbeit gefundenen Teilsequenzen sind wahrscheinlich den BAHD-Acyltransferasen in Klasse Va (Hydroxycinnamoyltransferasen, siehe Kapitel 1.4.1) zuzuordnen. Diese sind unter anderem an der Synthese von Chlorogensäure und Monolignolen beteiligt. Die hohe Anzahl von Genen zur Kodierung von BAHD-Acyltransferasen innerhalb der Pflanzengenome zeigt, dass es sich um eine Proteinsuperfamilie handelt. Immer neue Enzymaktivitäten, die den BAHD-Acyltransferasen zuzurechnen sind, werden beschrieben. 2006 besaß diese Proteinsuperfamilie noch keine 50 Mitglieder. Weiter wurden allein im Genom von *Arabidopsis* 61 BAHD-Gene identifiziert und in *Populus* sogar 94 BAHD-Gene (Yu et al., 2009).

Die Hydroxycinnamoyltransferase TpHCT2 aus *Trifolium pratense* katalysiert die Veresterung der Äpfelsäure (Sullivan, 2009), und eine weitere Hydroxycinnamoyltransferase aus *Arabidopsis thaliana* ist an der Biosynthese von Ligninintermediaten beteiligt (Hoffmann et al., 2003). Die Coniferylalkohol Acetyltransferase (CFAT) aus *Petunia hybrida* ist wichtig für die Biosynthese des

Blütenduftstoffs Ioeugenol (Dexter et al., 2007). Eine weitere interessante Hydroxycinnamoyltransferasen wurde aus *Hordeum vulgare* isoliert. Die Agmatin *N*-Hydroxycinnamoyltransferase ist hoch spezifisch für Agmatin als Acyl-Akzeptor und katalysiert die erste Reaktion der Biosynthese von Hydroxycinnamoylagmatin. Diese Substanz wird im Weiteren zum fungiziden Hordatol umgesetzt (Burhenne et al., 2003).

Weitere Beispiele könnten genannt werden, und je mehr BAHD-Acyltransferasen isoliert und biochemisch charakterisiert werden, desto besser ist deren Evolution nachzuvollziehen.

### **1.5. Die Hydroxycinnamoyltransferasen RAS, HST und HQT**

Die folgenden drei BAHD-Acyltransferasen gehören zur Klasse Va, wovon die ersten beiden Enzyme im Verlauf dieser Arbeit weiter untersucht werden:

- Hydroxycinnamoyl-CoA:Hydroxyphenyllactat Hydroxycinnamoyltransferase (Rosmarinsäuresynthase **(RAS)**)
- Hydroxycinnamoyl-CoA:Shikimat Hydroxycinnamoyltransferase **(HST)**
- Hydroxycinnamoyl-CoA:Chinat Hydroxycinnamoyltransferase **(HQT)**

Diese drei Enzyme sind CoA-Ester abhängige BAHD-Acyltransferasen, die an der Biosynthese von Rosmarinsäure (RAS), Chlorogensäure (HQT) und Monolignolen (HST) beteiligt sind. Sie übertragen 4-Cumaroyl-CoA, Caffeoyl-CoA und auch andere (Hydroxy-)Cinnamoyl-Reste aus den CoA-Thioestern auf ihre Akzeptorsubstrate (Petersen et al., 2009). RAS, HST und HQT setzen unterschiedliche Akzeptorsubstrate um. Die dabei gebildeten Produkte unterscheiden sich durch das Substitutionsmuster an den Aromaten.

Während die Grundzüge der Lignin-, Chlorogensäure- und Rosmarinsäure-Biosynthesewege aufgrund eingehender molekularbiologischer und biochemischer Untersuchungen weitgehend verstanden sind (siehe 1.3 Biosynthese der Rosmarinsäure), existieren kaum Informationen über die Evolution der beteiligten Gene bzw. Enzyme. Die Biosynthesewege weisen große Gemeinsamkeiten auf, und die beteiligten Enzyme scheinen sehr nah miteinander verwandt zu sein (Petersen et

al., 2009). Bei der Biosynthese von Caffeoylshikimat, Chlorogensäure und Rosmarinsäure wird 4-Cumaroyl-CoA aus dem allgemeinen Phenylpropanstoffwechsel als Hydroxyzimtsäuredonor verwendet (St. Pierre und De Luca, 2000).

Die Antwort auf die Frage, wie innerhalb des Phenylpropanstoffwechsels die Hydroxylierung der 4-Cumarsäure an Position 3 des aromatischen Ringes stattfindet, war lange Zeit unklar (siehe Kapitel 1.7 Biosynthese der Rosmarinsäure). Mittlerweile konnte gezeigt werden, dass diese Reaktion durch 4-Cumaroyl-3-Hydroxylasen (C3H) katalysiert wird. Diese Enzyme sind ähnlich der Zimtsäure 4-Hydroxylase (C4H) Cytochrom P450 abhängige Monooxygenasen (Franke et al., 2002).

Die Akzeptoren sind Derivate des Shikimatwegs: Shikimisäure (HST), Chinasäure (HQT) und 4-Hydroxyphenyllactat (RAS), das ausgehend von L-Tyrosin gebildet wird. Die Enzyme, die die Veresterungen katalysieren, sind Hydroxyzimtsäuretransferasen aus der Klasse Va der BAHD-Acyltransferase-Familie (siehe Kapitel 1.4 BAHD-Acyltransferasen) (Schoch et al., 2006). Die Hydroxyzimtsäuretransferasen zeigen hohe Sequenzähnlichkeiten untereinander und scheinen eine enge evolutionäre Beziehung zu haben. Trotzdem sind sie häufig sehr spezifisch für die Substrate ihres Biosynthesewegs. Für die Biosynthese von Rosmarinsäure im Speziellen mussten sehr wahrscheinlich nur wenige neue Enzymaktivitäten entwickelt werden. Sie sind möglicherweise aus Genen des Monolignol- oder Chlorogensäuremetabolismus hervorgegangen (Petersen et al., 2009).

Ein weiteres Beispiel für die nahe Verwandtschaft innerhalb der BAHD-Acyltransferasen sind Enzyme mit Hydroxycinnamoyl-CoA:Shikimat Hydroxycinnamoyltransferase (HST)- oder Hydroxycinnamoyl-CoA:Chinat Hydroxycinnamoyltransferase (HQT)-Aktivität. Durch die strukturelle Ähnlichkeit dieser beiden Enzyme wurden sogar HCSQTs beschrieben, die sowohl HST- als auch HQT-Aktivität besitzen und sowohl Shikimat als auch Chinat als Akzeptorsubstrate verwenden. Meistens liegt jedoch eine Präferenz für eines der beiden Akzeptorsubstrate vor. Dies zeigt, dass in einigen Fällen die Änderung weniger Aminosäuren ausreicht, um die Substratspezifität der Enzyme zu variieren (Sander und Petersen, 2011).

## 1.6. Rosmarinsäure

Rosmarinsäure ist ein Ester aus Kaffeesäure und 3,4-Dihydroxyphenylmilchsäure (Abb. 1.1). Im Jahr 1958 wurde zum ersten Mal Rosmarinsäure aus *Rosmarinus officinalis* isoliert (Scarpati und Oriente, 1958).

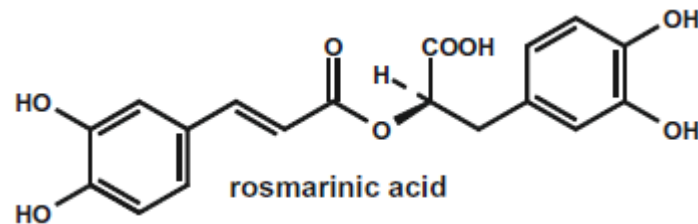


Abb.1.1: Chemische Struktur der Rosmarinsäure, die ein Ester aus Kaffeesäure und 3,4-Dihydroxyphenylmilchsäure ist (Formel aus Petersen et al., 2009).

Rosmarinsäure macht den Hauptbestandteil der sogenannten Labiatengerbstoffe aus. Sie besitzt antivirale, antibakterielle und antiinflammatorische Eigenschaften (Petersen und Simmonds, 2003). Die antiinflammatorischen Eigenschaften von Rosmarinsäure sind auf die Inhibition der C3-Konvertase des Komplementsystems zurückzuführen (Englberger et al., 1988). Weiterhin hemmt sie zusätzlich die Prostaglandinsynthese (Rampart et al., 1986). Rosmarinsäure wird jedoch noch nicht als reines medizinisches Produkt verwendet.

Viele Pflanzen aus der Familie der Lamiaceae finden u.a. wegen ihres Rosmarinsäuregehalts breite Anwendungen. Dabei werden die reduzierenden und radikalfangenden Eigenschaften von Rosmarinsäure genutzt. Weiterhin sind in Deutschland Fertigarzneimittel zugelassen, deren Wirksamkeit u.a. auf den Rosmarinsäure-Gehalt zurückzuführen ist:

- Lomaherpan<sup>®</sup> - Creme gegen Lippenherpes, enthält Melissenextrakt
- Traumaplant<sup>®</sup> - Salbe gegen Muskelschmerzen und Sportverletzungen, enthält Beinwelleextrakt

- Allerpril® - Nahrungsergänzungsmittel gegen Allergien, enthält Öl aus *Perilla frutescens*
- Salvysat® - Hemmung der Schweiß- und Milchdrüsensekretion, äußerlich im Mund-, Rachenraum, antibakterielle Effekte und entzündungshemmend, enthält Inhaltsstoffe aus Salbei

## 1.7. Biosynthese der Rosmarinsäure

An der Biosynthese der Rosmarinsäure sind folgende acht Enzyme des Sekundärstoffwechsels beteiligt (siehe Abb. 1.2.):

- Phenylalanin Ammoniak-Lyase (PAL)
- Zimtsäure 4-Hydroxylase (C4H)
- Hydroxyzimtsäure:Coenzym A-Ligase (4CL)
- Tyrosin Aminotransferase (TAT)
- Hydroxyphenylpyruvat Reduktase (HPPR)
- Hydroxycinnamoyl-CoA:Hydroxyphenyllactat Hydroxycinnamoyltransferase = Rosmarinsäure Synthase (RAS)
- 4-Cumaroyl-4'-Hydroxyphenyllactat 3/3'-Hydroxylasen (4C-pHPL 3H bzw. 4C-pHPL 3'H)
- Caffeoyl-4'-Hydroxyphenyllactat 3'-Hydroxylase (Caf-pHPL 3'H)

Die genannten Enzyme konnten mit Zellkulturen von *Anchusa officinalis*, *Coleus blumei* und *Melissa officinalis* aufgeklärt werden (De-Eknamkul und Ellis, 1987; Petersen et al., 1993; Weitzel und Petersen, 2011). Die beiden Aminosäuren L-Phenylalanin und L-Tyrosin dienen als Ausgangssubstanzen für die Biosynthese der Rosmarinsäure (Ellis und Towers, 1970) (siehe Abb. 1.2.).

Im Phenylpropanstoffwechsel wird 4-Cumaroyl-CoA als energiereiches Intermediat aus L-Phenylalanin gebildet, das den Ausgangspunkt für die Synthese von Rosmarinsäure darstellt. L-Phenylalanin wird dazu in zwei enzymkatalysierten Schritten (Phenylalanin Ammoniak-Lyase (PAL), Zimtsäure 4-Hydroxylase (C4H))



desaminiert und hydroxyliert (Razzaque und Ellis, 1977). Danach erfolgt die Aktivierung (Hydroxyzimtsäure:Coenzym A-Ligase (4CL)) zu 4-Cumaroyl-CoA (Karwatzki et al., 1989).

Die zweite Aminosäure, L-Tyrosin, wird zu 4-Hydroxyphenylpyruvat transaminiert (Tyrosin Aminotransferase (TAT)) (De Eknankul und Ellis, 1987; Mizukami und Ellis, 1991; Kempin, 1994). Anschließend wird 4-Hydroxyphenylpyruvat zu *R*(+)-4-Hydroxyphenyllactat reduziert (Hydroxyphenylpyruvat Reduktase (HPPR)) (Petersen und Alfermann, 1988; Häusler et al., 1991; Kim et al., 2004).

Die weiteren enzymatischen Reaktionen sind spezifisch für die Synthese von Rosmarinsäure (siehe Abb. 1.2.):

Die aus L-Phenylalanin und L-Tyrosin entstandenen Produkte 4-Cumaroyl-CoA und 4-Hydroxyphenyllactat werden zu 4-Cumaroyl-4'-Hydroxyphenyllactat verestert (Petersen und Alfermann, 1988; Petersen, 1991; Petersen et al., 1993). Diese Veresterung wird durch die Rosmarinsäure Synthase (Hydroxycinnamoyl-CoA:Hydroxyphenyllactat Hydroxycinnamoyltransferase (RAS)) katalysiert. Die RAS wurde erstmals 1988 aus Zellkulturen von *Coleus blumei* isoliert und charakterisiert (Petersen und Alfermann, 1988). Als natürliche Substrate dieses Enzyms wurden 4-Cumaroyl-CoA und 4-Hydroxyphenyllactat ermittelt. Vollängen-cDNAs der RAS konnten aus *Coleus blumei*, *Melissa officinalis* und *Glechoma hederacea* isoliert, in *E. coli* exprimiert und biochemisch charakterisiert werden (Berger et al., 2006; Sander, 2010; Weitzel und Petersen, 2011).

Zum Abschluss wird der entstandene Ester durch zwei Cytochrom P450-abhängige Enzymaktivitäten zur Rosmarinsäure hydroxyliert (4-Cumaroyl-4'-Hydroxyphenyllactat 3/3'-Hydroxylasen (4C-pHPL 3H bzw. 4C-pHPL 3'H) und Caffeoil-4'-Hydroxyphenyllactat 3'-Hydroxylase (Caf-pHPL 3'H)) (Petersen, 1997; Eberle et al., 2008).

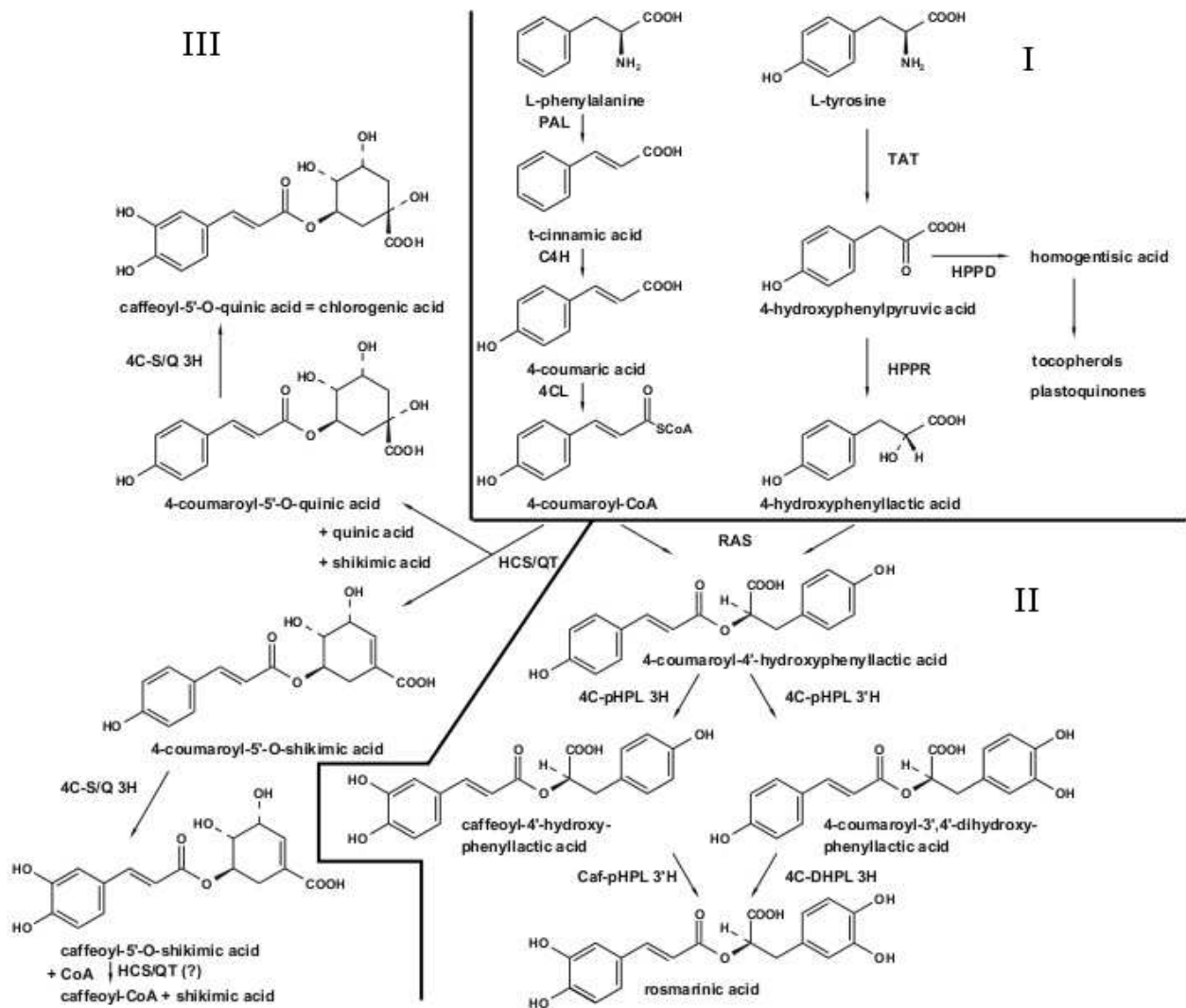


Abbildung 1.2: Biosynthesewege der Rosmarinsäure, Chlorogensäure und Caffeoylshikimisäure (nach Petersen et al., 2009). Im allgemeinen Phenylpropanstoffwechsel (I) sind folgende Enzyme aktiv: Phenylalanin Ammoniak-Lyase (PAL); Zimtsäure 4-Hydroxylase (C4H); Hydroxyzimtsäure:CoA-Ligase (4CL). Tyrosin Aminotransferase (TAT) und Hydroxyphenylpyruvat Reduktase (HPPR) werden für die Bereitstellung des 4-Hydroxyphenyllactats benötigt. Die Synthese von Rosmarinsäure (II) wird durch folgende Enzyme katalysiert: Rosmarinsäure Synthase (Hydroxycinnamoyl-CoA:Hydroxyphenyllactat Hydroxycinnamoyltransferase (RAS)); 4-Cumaroyl-4'-Hydroxyphenyllactat 3/3'-Hydroxylasen (4C-pHPL 3H, 4C-pHPL 3'H) oder Caffeoyl-4'-Hydroxyphenyllactat 3'-Hydroxylase (Caf-pHPL 3'H); 4-Cumaroyl-3',4'-Dihydroxyphenyllactat 3-Hydroxylase (4C-DHPL 3H).

Die Synthese von Chlorogensäure und der Monolignol-Vorstufe Caffeoyleshikimisäure (III) zeigen große Ähnlichkeiten zur Synthese der Rosmarinsäure (II). An der Synthese sind die Enzyme Hydroxycinnamoyl-CoA:Shikimat Hydroxycinnamoyltransferase (HST) oder Hydroxycinnamoyl-CoA:Chinat Hydroxycinnamoyltransferase (HQT) beteiligt. Durch die strukturelle Ähnlichkeit dieser beiden Enzyme wird in der Abbildung die Abkürzung HCS/QT verwendet. HCS/QTs sind Enzyme, die sowohl HST- als auch HQT-Aktivität besitzen, wobei meistens eine Präferenz für eine der Reaktionen vorliegt.

### **1.8. Vorkommen der Rosmarinsäure im Pflanzenreich**

Rosmarinsäure ist im Pflanzenreich weit verbreitet. In der Familie der Boraginaceae ist sie durchgängig zu finden, wohingegen sie bei den Lamiaceae weitgehend auf die Unterfamilie der Nepetoideae beschränkt ist (Litvinenko et al., 1975). Weiterhin kommt Rosmarinsäure in verschiedenen anderen Pflanzentaxa vor, wie z.B. bei Hornmoosen (Anthocerotaceae; Takeda et al., 1990), Farnen der Gattung *Blechnum* (Blechnaceae; Harborne, 1965) oder bei Seegräsern (Zosteraceae; Mølgaard und Ravn, 1988). In Tab. 1.1. sind die bis jetzt bekannten Ergebnisse des Vorkommens von Rosmarinsäure im Pflanzenreich erfasst.

Tab. 1.1: Aufzählung der bekannten Rosmarinsäure produzierenden Pflanzenfamilien (aus Petersen et al., 2009)

|   |                |  |
|---|----------------|--|
| <i>Anthocerophytina</i>                     | Anthocerotales | Anthocerotaceae  |
| <i>Filicophytina</i>                        | Blechnales     | Blechnaceae  |
| <i>Spermatophytina, class Magnoliopsida</i> |                |  |
| Basal orders                                | Chloranthales  | Chloranthaceae   |
| Monocoty ledonous plants                    | Alismatales    | Araceae<br>Potamogetonaceae<br>Zosteraceae                     |
|   | Zingiberales   | Cannaceae<br>Marantaceae                                       |
| Eudicoty ledons                             | Liliales       | Melanthaceae   |
|   | Myrtales       | Onagraceae   |
|   | Celastrales    | Celastraceae   |
|   | Rosales        | Rosaceae   |
|   | Cucurbitales   | Cucurbitaceae  |
|   | Malvales       | Malvaceae/Sterculiaceae<br>Malvaceae/Tiliaceae                 |
|   | Gentianales    | Rubiaceae  |
|   | Lamiales       | Lamiaceae<br>Plantaginaceae<br>Acanthaceae<br>Scrophulariaceae |
|   | Unclear        | Boraginaceae/<br>Hydrophyllaceae                               |
|   | Apiales        | Apiaceae   |
|   |                | Araliaceae   |
|   | Asterales      | Asteraceae   |
|   | Dipsacales     | Dipsacaceae  |

Hornmoose (Anthocerotaceae) sind eine der frühesten Landpflanzenarten, die sich wahrscheinlich aus einer Familie der Grünalgen (Charophyceae) entwickelt haben. Rosmarinsäure konnte auch in Gattungen dieser frühen Landpflanzen nachgewiesen werden (Takeda et al., 1990). Die Möglichkeit Rosmarinsäure, sowie Chlorogensäure und Monolignole zu bilden, ist somit eine sehr alte Eigenschaft, die weit verstreut im Reich der Pflanzen zu finden ist.

## **1.9. Phylogenie der Lamiaceae**

Aufgrund der Blütenstruktur gehören Lamiaceae (auch Labiatae, Lippenblüter) zu einer der höher entwickelten Pflanzenfamilien der Echten Eudikotyledonen. Zu den Lamiaceae zählen über 7000 Species in 230 Gattungen (Harley et al., 2004). Die Blätter sind häufig gegenständig, dekussiert angeordnet und der Stängel meist vierkantig geformt. Vorwiegend sind Kräuter oder Sträucher vertreten. Die arttypische Infloreszenz ist ein Thyrsus mit cymösen Teilblütenständen. Die Blüten bilden dabei einen zygomorphen zweilippigen Körper mit vierteilig gespaltenem Fruchtknoten, dessen Samenanlagen laminal oder submarginal plazentiert sind. Die variablen Strukturen der genannten morphologischen Merkmale innerhalb der Familie der Lamiaceae lassen eine eindeutige Klassifikation jedoch nicht zu (Hedge, 1992).

Die Lamiaceae wurden lange Zeit als nahe Verwandte der Verbenaceae angesehen, jedoch war deren genaue Abgrenzung nicht geklärt. Dadurch kam es zu falschen Einordnungen verschiedener Unterfamilien. Schon van Tieghem (1907) und Junell (1934) erkannten, dass die Stammbäume der Lamiaceae und Verbenaceae nicht eindeutig voneinander abgegrenzt waren. 1992 wurde diese Vermutung durch Cantino et al. bestätigt, die zeigen konnten, dass der damalige Aufbau der Lamiaceae polyphyletisch war. Eine natürlichere Klassifizierung konnte erzielt werden, indem die Unterfamilien der Caryopteridoideae, Chloanthoideae, Viticoideae und die Gattung Monochileae weiterhin zu den Lamiaceae gezählt werden und nur die Verbenoideae mit charakteristischen razemösen Infloreszenzen innerhalb der Verbenaceae verblieben. Diese Klassifizierung konnte auch durch spätere cpDNA-Analysen unterstützt werden (Kadereit, 2004).

Historisch erfolgte die Einordnung von neu entdeckten Pflanzen in die verschiedenen Gattungen der Lamiaceae vor Bentham (1836) eher willkürlich (Lindley, 1829). Die taxonomische Gliederung der Lamiaceae von Bentham (1836) konnte erstmals einen strukturierten Aufbau präsentieren, der eine stringente Einteilung der zuvor uneinheitlichen Gattungen aufzeigte. Bentham (1836) gliederte die Familie der Lamiaceae in elf Unterfamilien. In späteren Bearbeitungen (Bentham, 1876) wurde die Anzahl der Unterfamilien auf acht reduziert (Bräuchler, 2009).

## **Unterfamilien der Lamiaceae**

*Sensu* Bentham (1836 und 1876)

Ocimoideae

Menthoideae (1876 gestrichen)

Monardeae

Satureineae

Melissineae (1876 gestrichen)

Scutellarineae (1876 gestrichen)

Prostanthereae

Nepeteae

Stachydeae

Prasieae

Ajugoideae

Weitere an die Gliederung von Bentham angelehnte Konzepte zur Gliederung wurden z.B. von Briquet (1897) präsentiert. Ein radikaler Schnitt wurde 1945 von Erdtman publiziert, welcher das Verständnis der Taxonomie der Lamiaceae grundlegend änderte. Erdtmann (1945) teilte das Reich der Lamiaceae in die übergeordneten Familien Lamioideae und Nepetoideae auf. Als Unterscheidungsmerkmal dienen hierbei die Pollen:

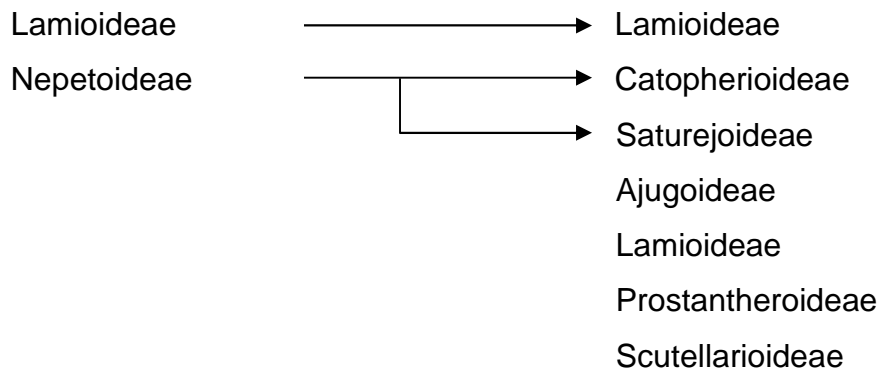
- Lamioideae besitzen tri- oder tetracolpate, zweikernige Pollen
- Nepetoideae besitzen hexacolpate, dreikernige Pollen.

Die Einteilung von Erdtman (1945) wurde 1967 von Wunderlich übernommen und erweitert. Zusätzlich zu den palynologischen Merkmalen wurde auch die Samenanatomie mit einbezogen. Als Ergebnis dieser zusammenhängenden Betrachtung wurde eine weitere Unterteilung auf sechs Unterfamilien von Wunderlich vorgenommen. Die Unterfamilie Lamioideae wurde beibehalten, Nepetoideae aufgeteilt und vier weitere Unterfamilien angefügt:

## Unterfamilien der Lamiaceae

*Sensu* Erdtman (1945)

*Sensu* Wunderlich (1967)



Eine der letzten Arbeiten, die sich rein auf eine Definition der Unterfamilien der Lamiaceae durch morphologische Daten bezieht, wurde 1992 von Cantino et al. präsentiert. Durch kladistische Analysen wurde eine Einteilung vorgenommen, die von den vorherigen Definitionen abwich und eine Gliederung in acht Unterfamilien vorsah:

## Unterfamilien der Lamiaceae

*Sensu* Cantino et al. (1992)

Ajugoideae  
Chloanthoideae  
Lamioideae  
Nepetoideae  
Pogostemonoideae  
Scutellarioideae  
Teucrioideae  
Viticoideae

Von hier an wird die Unterteilung durch molekular-kladistische Untersuchungen erweitert. Die bisherigen taxonomischen Konzepte werden durch molekularbiologische Eigenschaften überprüft und wenn nötig korrigiert oder ergänzt. Eine Überprüfung des taxonomischen Konzepts nach Cantino et al. (1992) wurde von Wagstaff et al. (1998) vorgenommen. Durch Sequenzvergleiche des *rbcL*-

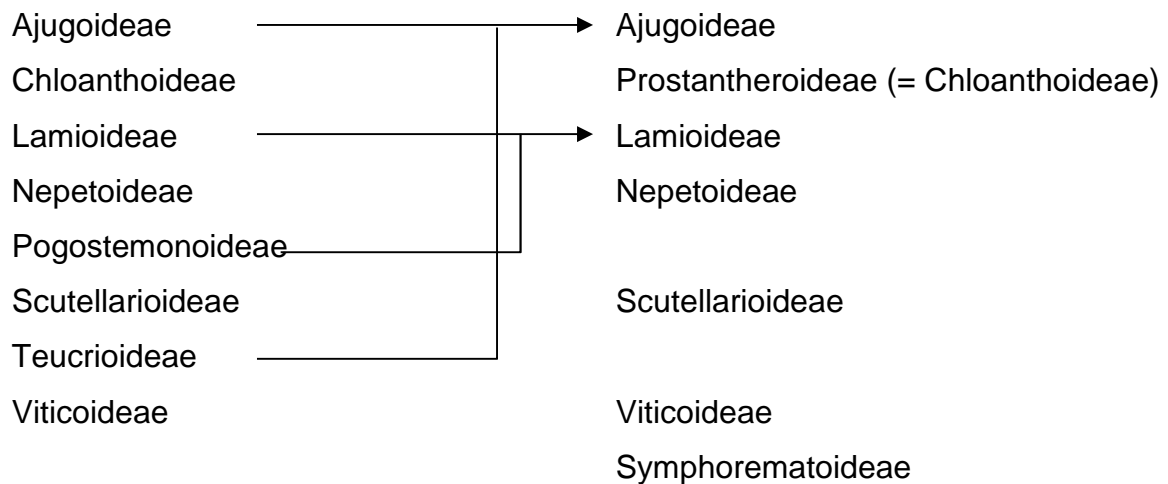
(chloroplast **Rubisco** large subunit) und des *ndhF*-Gens (NADH dehydrogenase) wird die Einteilung von Teucrioideae, Scutellarioideae, Lamioideae, Pogostemonoideae und Nepetoideae als natürliche Gruppen bestätigt.

Nach einer neueren Gliederung, die verschiedene molekularbiologische Analysen berücksichtigt, werden einige Änderungen der Gliederung (Cantino et al., 1992) in Bezug auf Namensgebung und Unterteilung vorgenommen (Harley et al., 2004). Es werden sieben Unterfamilien anerkannt:

### Unterfamilien der Lamiaceae

*Sensu Cantino et al. (1945)*

*Sensu Harley et al. (2004)*



Wesentliche Unterschiede gegenüber der Einteilung nach Cantino et al. (1992) sind die Zusammenfassung von Ajugoideae und Teucrioideae, sowie Lamioideae und Pogostemonoideae. Chloanthoideae werden in Prostantheroideae umbenannt. Die Symphoremataceae werden als neue Unterfamilie in die Lamiaceae mit einbezogen. Für die großen Unterfamilien, wie Lamioideae und Nepetoideae, deren Monophylie gut belegt ist, sind voraussichtlich keine größeren Änderungen mehr zu erwarten (Harley et al., 2004; Scheen et al., 2010). Die aktuelle Klassifikation der Angiosperm Phylogeny Group (APG Version III) übernimmt die Einteilung mit sieben Unterfamilien nach Harley et al. (2004) (siehe <http://www.mobot.org/MOBOT/research/APweb/>).

Die Nepetoideae sind mit fast 3800 eingeordneten Arten die artenreichste und die am klarsten abgegrenzte Unterfamilie innerhalb der Lamiaceae. Die meisten in dieser Arbeit analysierten Pflanzen gehören zu dieser Unterfamilie. Nach Harley et



al. (2004) können die Nepetoideae in die drei Tribus Mentheae, Elsholtzieae und Ocimeae unterteilt werden. Alle bisher durchgeführten genetischen Untersuchungen bestätigten, dass die Unterfamilie Nepetoideae monophyletisch ist (Wagstaff et al., 1995). Dies wurde bereits durch morphologische Merkmale unterstützt wie z.B. den hexacolpaten dreikernigen Pollen (Erdtman, 1945). Bei der jungen Pflanze ist die Basis der Kotyledonen verlängert. Dadurch sind das Hypokotyl und die Radicula größtenteils überdeckt. Ein weiteres typisches charakteristisches Merkmal der Nepetoideae sind verhältnismäßig hohe Konzentrationen an leicht flüchtigen Monoterpenen. Diese sind häufig abgesondert in spezialisierten Drüsen oder Trichomen lokalisiert (Wink, 2003). Einige Arten produzieren zusätzlich Sesquiterpene. Außerdem produzieren Nepetoideae phenolische Substanzen, hauptsächlich repräsentiert durch die Rosmarinsäure.

Generelle typische Sekundärmetabolite der Lamiaceae sind Terpenoide, wie Mono-, Sesqui-, Di- und Triterpene. Bei Terpenen handelt es sich um reine Kohlenwasserstoffe, während Terpenoide funktionelle Gruppen enthalten. Als Grundbaustein für Terpenoide dient Isopren. Auch phenolische Substanzen, vor allem Phenolsäuren, wie die genannte Rosmarinsäure oder Flavonoide sind stark vertreten. Stickstoffhaltige Sekundärmetabolite spielen eine eher untergeordnete Rolle (Hegnauer, 1989).

### **1.10. Phylogenie der Boraginaceae**

Über einen langen Zeitraum wurde vermutet, dass die Familien der Lamiaceae und der Boraginaceae (Raublattgewächse oder Borretschgewächse) nahe verwandt sind. Jedoch wurden durch molekularbiologische Untersuchungen der *rbcL*-Gensequenzen festgestellt, dass die Familie der Boraginaceae eher der Familie der Solanaceae nahe steht (Grayer et al., 1999). Die aktuelle Klassifikation wurde 2009 von der Angiosperm Phylogeny Group (APG) veröffentlicht (Version III). In dieser letzten Version wurden die Boraginaceae ohne Zuweisung in eine Ordnung in die Unterklasse der Lamiidae eingesetzt (Chase und Reveal, 2009).

Die Boraginaceae umfassen 148 bis 156 Gattungen mit 2740 Arten, die weltweit vertreten und von denen die meisten ein- bis mehrjährige krautige Pflanzen,

Sträucher und Bäume sind. Die Boraginaceae produzieren u.a. Pyrrolizidinalkaloide als Sekundärstoffe gegen Fraßfeinde (Hartmann, 2004).

Die Boraginaceae wurden anhand molekularphylogenetischer Daten (APG III, 2009) in sechs Unterfamilien eingeteilt, wobei beispielsweise die Unterfamilie Hydrophylloideae bis vor kurzem als eigenständige Familie Hydrophyllaceae geführt wurde:

- Boraginoideae
- Cordioideae
- Ehretioideae
- Heliotropoideae,
- Hydrophylloideae
- Lennooideae

Die Boraginoideae sind vorwiegend krautige Pflanzen, wogegen in den anderen Unterfamilien auch baumförmige Vertreter vorkommen (Bresinsky et al., 2008).

### **1.11. Zielsetzung**

Die Hydroxycinnamoyltransferase RAS ist ein spezifisches Enzym der Rosmarinsäurebiosynthese und gehört zur Proteinsuperfamilie der BAHD-Acyltransferasen, die im Pflanzenreich ubiquitär verbreitet sind. Daneben existieren HCTs, die spezifische Enzyme der Chlorogensäure- und Monolignol-Biosynthese sind. Diese katalysieren sehr ähnliche Reaktionen wie die RAS und weisen eine recht hohe Sequenzähnlichkeit zur RAS auf. Bislang deuten verfügbare Ergebnisse darauf hin, dass RAS, HST und HQT sehr nah miteinander verwandt sind (Petersen et al., 2009).

Anhand der Sequenzähnlichkeit kann die Abstammungsgeschichte von Hydroxycinnamoyltransferasen der BAHD-Acyltransferase-Familie rekonstruiert werden. Je mehr Sequenzen aus verschiedenen Arten miteinander verglichen werden, desto genauer sind die Ergebnisse. Durch Vergleich von bisher bekannten und in dieser Arbeit neu gefundenen HCT-Genen aus den verschiedensten

Pflanzenarten sollen in dieser Arbeit insbesondere Hinweise auf die evolutionäre Entwicklung der Rosmarinsäuresynthese gewonnen werden. Die Frage der mono- oder polyphyletischen Entwicklung der Rosmarinsäuresynthese in weit voneinander entfernt stehenden Taxa soll näher untersucht werden.

In Pflanzenarten, die nachweislich Rosmarinsäure produzieren, soll nach RAS-Genen gesucht werden. Dazu werden anhand der konservierten Motive der BAHD-Acyltransferasen und den bislang bekannten RAS-, HST- und HQT-Sequenzen degenerierte Primer abgeleitet. Werden RAS-Sequenzen gefunden, ist davon auszugehen, dass in dieser Pflanzenart die Rosmarinsäurebiosynthese über dieselben Reaktionsschritte abläuft, wie es in den Lamiaceen *Coleus blumei* (Petersen et al., 1993) und *Melissa officinalis* (Weitzel und Petersen, 2011) nachgewiesen werden konnte.

Besonders interessant ist, ob in evolutionär weit auseinander liegenden Pflanzenarten RAS-Sequenzen zu finden sind. Wäre dies der Fall, so könnte bei einer hohen Sequenzidentität zwischen RAS-Sequenzen davon ausgegangen werden, dass die Fähigkeit, Rosmarinsäure zu produzieren eine alte Eigenschaft ist. Viele Arten hätten somit im Laufe der Evolution diese Fähigkeit wieder abgelegt. Existieren jedoch Enzyme mit RAS-Aktivität, aber kaum oder keiner Sequenzähnlichkeit zu den bislang bekannten RAS-Genen, so hat sich diese Fähigkeit womöglich mehrfach unabhängig voneinander entwickelt.

## 2. Material

### 2.1. Geräte

|                                   |  |                                    |
|-----------------------------------|--|------------------------------------|
| Autoklav                          | Systec VX-150<br>Advantage –Lab AL02-02-100    | Systec GmbH<br>Labor-Systemtechnik |
| DC-Platten                        | TLC Silica gel 60 F <sub>254</sub>             | Merck                              |
| DC-Platten                        | TLC Cellulose F                                | Merck                              |
| Dialysierschlauch                 | ZelluTrans                                     | Roth                               |
| Elektrophoresesystem für SDS-PAGE | PerfectBlue™ Twin                              | PEQLAB<br>Biotechnologie GmbH      |
| Elektroporationsküvette           | Electroporation Cuvettes 1 mm                  | Eppendorf                          |
| Elektroporator                    | MicroPulser™                                   | Bio-Rad                            |
| Entsalzungssäulen                 | PD-10 columns                                  | GE Healthcare                      |
| Gefriertrockner                   | Christ L1                                      | B. Braun Biotech                   |
| Geldokumentationssystem           | Quick Store „plus“                             | RS-Laborgeräte                     |
| HPLC                              | SpectroMonitor® 3200                           | LDC Analytical                     |
| HPLC                              | L-6000 Pump                                    | Merck-Hitachi                      |
| HPLC                              | D-2500 Chromato-Integrator                     | Merck-Hitachi                      |
| Magnetrührer                      | MR 3001  | Heidolph Instr.                    |
| PCR-Cycler                        | RoboCycler® Gradient 96<br>Biorad<br>Eppendorf | Stratagene<br>Eppendorf, Biorad    |
| pH-Meter                          | Accumed Basic                                  | Fisher Scientific                  |
| Photometer                        | Eppendorf BioPhotometer                        | Eppendorf                          |
| Photometer                        | UVIKON Spectrophotometer 930                   | Kontron Instruments                |
| Plattformschüttler                | Duomax 1030                                    | Heidolph                           |
| Rotationsverdampfer               | Rotavapor RE 120                               | Büchi                              |
| Säulenmaterial                    | Ni-NTA His Bind® Resin                         | Novagen                            |

|                                 |                                    |                                |
|---------------------------------|------------------------------------|--------------------------------|
| Schüttler für Bakterienkulturen | -                                  | Gallenkamp                     |
| Speed Vac                       | UNIVAPO 100 H                      | UniEquip                       |
| Sterilbank                      | Gelaire Laminar Air Flow Class 100 | Gelman Instrument              |
| Sterilfiltration                | Rotilabo® Spritzenfilter steril    | Roth                           |
| Thermomixer                     | Thermomixer Comfort                | Eppendorf                      |
| Ultraschallbad                  | Sonorex Super RK 510 H             | Bandelin                       |
| Ultraschallprozessor            | UP200s                             | Dr. Hielscher GmbH             |
| UV-Lampe                        | HL-6-KM                            | Bachofer GmbH                  |
| Vortex                          | Vortex-Genie 2                     | Scientific Industries          |
| Waagen                          | PT 310                             | Sartorius                      |
| Waagen                          | Analysenwaage EG 300-3M            | Kern                           |
| Waagen                          | Analysenwaage H64                  | Mettler                        |
| Wasseraufbereitung              | Milli-RX 20                        | Millipore                      |
| Zentrifugen                     | Centrifuge 5415D                   | Eppendorf                      |
| Zentrifugen                     | Sigma 3K30                         | B.Braun Biotech. International |
| Zentrifugen                     | Biofuge 17RS                       | Heraeus Sepatech               |

## 2.2. Chemische Grundsubstanzen

|                                    |                        |
|------------------------------------|------------------------|
| (Hydroxy)Zimtsäure-Succinimidester | Sammlung des Instituts |
| (N)-Lauroyl-Sarcosin, Na-Salz      | Sigma                  |
| 1,4-Dithiothreitol                 | Serva                  |
| Aceton                             | Roth                   |
| Acrylamid: rotiphoese Gel 30       | Roth                   |
| Agar-Agar                          | C.E. Roeper GmbH       |
| Agarose NEEQ, Ultra Qualität       | Roth                   |

|                                     |                        |
|-------------------------------------|------------------------|
| Ammoniak 25%                        | Riedel-de-Haen         |
| Ampicillin, Na-Salz                 | Roth                   |
| Cetyltrimethylammoniumbromid (CTAB) | Roth                   |
| Chloroform                          | Roth                   |
| Chlorogensäure (Hemihydrat)         | Fluka                  |
| Coenzym A-Trilithium-Salz           | AppliChem              |
| Coomassie Brilliant Blue G250       | Fluka                  |
| Coomassie® Brillant Blue R250       | Fluka                  |
| Ethidiumbromid-Lösung 1%            | AppliChem              |
| Guanidinthiocyanat                  | Roth                   |
| Imidazol                            | Roth                   |
| L-(+)-Ascorbinsäure                 | Roth                   |
| Methanol                            | Fisher Scientific      |
| Natriumcitrat                       | Sigma                  |
| Natriumdihydrogenphosphat           | Duchefa                |
| Polyvinylpyrrolidon MW 40.000       | Sigma                  |
| Rosmarinsäure                       | Sammlung des Instituts |
| Streptomycin                        | Roth                   |
| Triton X-100                        | Sigma                  |
| Zimtsäure                           | Duchefa Biochemie      |

Alle weiteren Chemikalien wurden von den Firmen Fluka, Roth, Sigma und Merck in den Qualitäten p.a. oder reinst bezogen. Für sämtliche Arbeiten wurde das entmineralisierte Wasser der hauseigenen Herstellung verwendet. Für molekularbiologische Arbeiten wurde das Wasser zusätzlich autoklaviert.

## 2.3. Vektorkarten

### pGEM®-T-Vektor

Fa. Promega

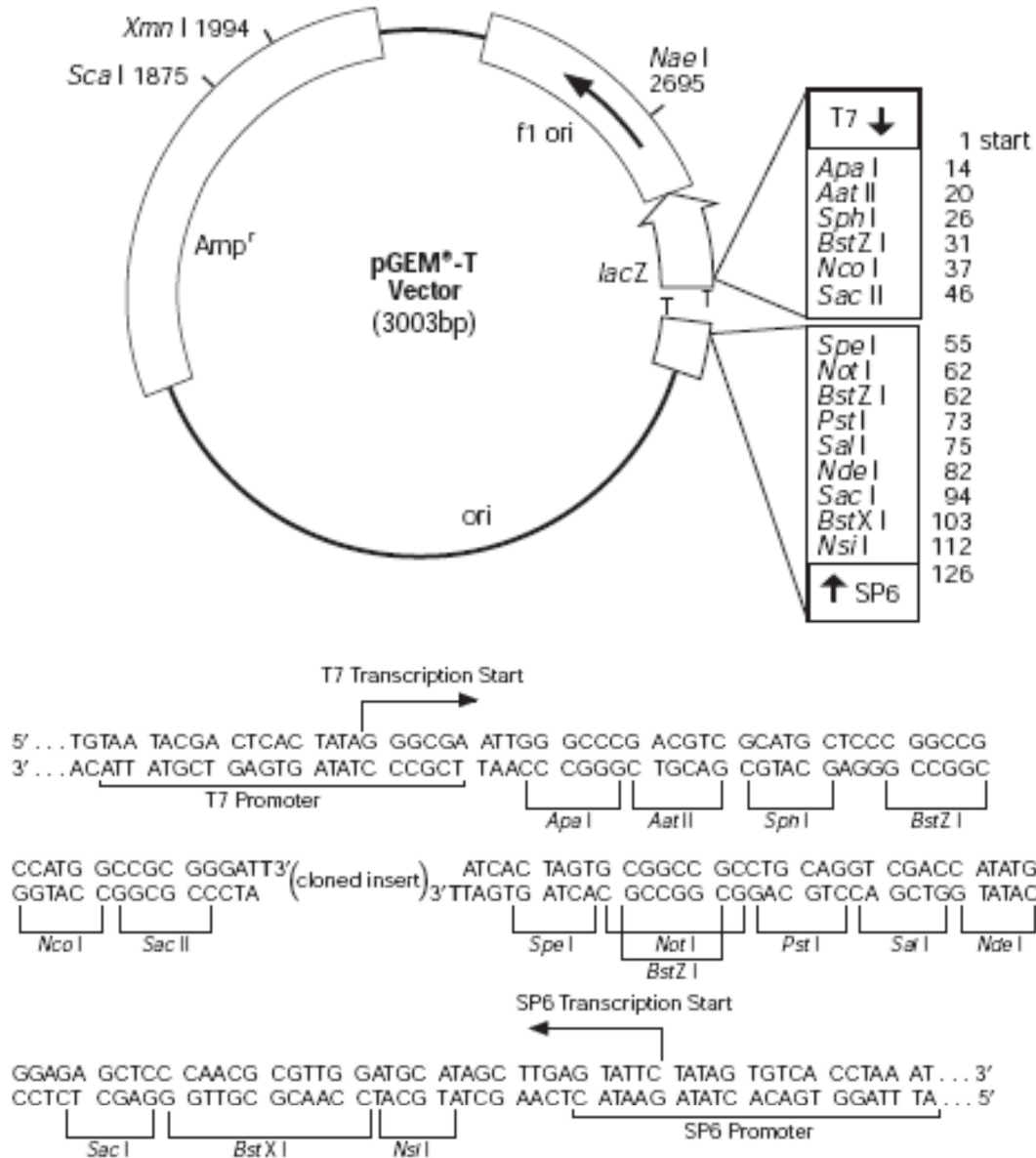


Abb. 2.1: Klonierungsvektor pGEM®-T und zugehörige Multiple Cloning Site (Quelle: <http://www.promega.com/~media/files/resources/protocols/technical%20manuals/0/pgem-t%20and%20pgem-t%20easy%20vector%20systems%20protocol.pdf>)

## pET-15b-Vektor

Fa. Novagen

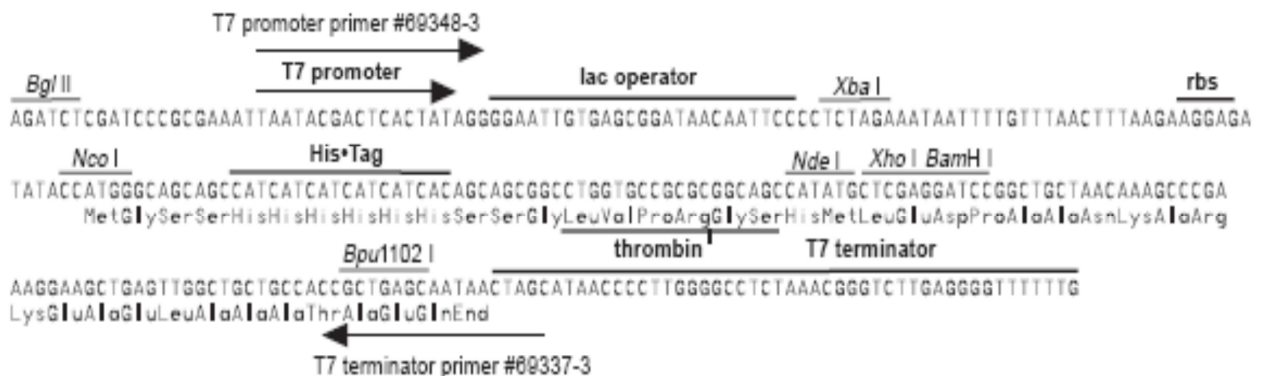
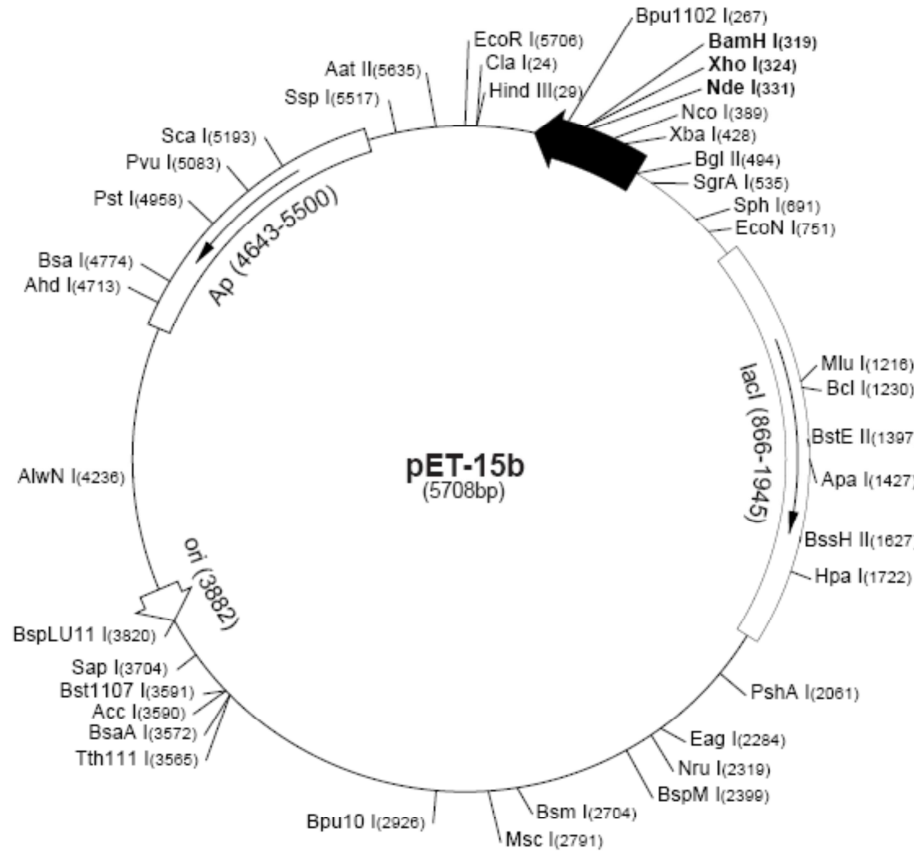


Abb. 2.2: Expressionsvektor pET-15b und zugehörige Multiple Cloning Site (Quelle: <http://www.addgene.org/vector-database/2543/>)



## **2.4. Primer**

### **MP-RAS-VL5 Forward**

5'-ATTACATATGAAGATAGAAGTCAAAGACTC-3'

### **MP-RAS-VL3 Reverse**

5'-TAGGATCCTCATCAAATCTCATAAAACAACTTCTC -3'

### **RASgen Forward**

5'-GATGAKYTYGGYGAYTT-3'

### **RAS DFGWG Reverse**

5'-CCCCANCCRAARTC-3'

### **RAS-Fn Forward**

5'-CCGCCCGGAGCTCAAC-3'

### **RAS-Rn Reverse**

5'-CCGGCTCACGCCGTCG-3'

### **RAS-Fn2 Forward**

5'-CGCCCCTACCCCGCTGCCGC-3'

### **RAS-Rn2 Reverse**

5'-CTGTGAGCTCCCTGCGCCAGCTTGC-3'

### **RASgen-Fn3 Forward**

5'-GGCGCCGGCAAATCC-3'

### **RASgen-Rn3 Reverse**

5'-GGATTTGCCGGCGCC-3'

### **RAS-FN7 Forward**

5'-GCAAACGAGCACC ACTTATCCGACGGCGTCGCCGCC-3'

### **RAS-RN6 Reverse**

5'-CACCGCCTTCCCCCACC(AG)AA(AG)TCCGGCTC(AG)TA-3'

### **RAS-VL5-SO Forward**

5'-ATTACATATGGAAATGAAAACCAATGG-3'

### **VL5-SO Forward**

5'-ATTACATATGATGAAGATAGATGTGCAAG-3'

### **VL3-SO Reverse**

5'-TAGGATCCTCATCATATATCATAAAAGATCTTCTG-3'

### **HST700f Forward**

5'-GGVATGRCARCA YCAYGYDGCMGAYGGATTTTCYGGBCTT-3'

**HST700r**                      **Reverse**  
5`-AAKATMGGMCKMCCCCVCCAAARTCWGCATCAT-3`

## 2.5. *Molekularbiologische Reagenzien und Kits*

|  |                |
|--|----------------|
| GeneRacer™ Core Kit                        | Invitrogen     |
| GeneRacer™ SuperScript III RT Module       | Invitrogen     |
| GoTaq® Flexi DNA Polymerase                | Promega        |
| High Fidelity PCR Enzyme Mix               | Fermentas      |
| NucleoSpin® Extract II                     | Macherey-Nagel |
| pET-15b                                    | Novagen        |
| pGEM®-T-Vector System I                    | Promega        |
| QIAprep Spin Miniprep Kit                  | QIAGEN         |
| RevertAid™ First Strand cDNA Synthesis Kit | Fermentas      |
| Roti®-Transform                            | Roth           |
| GeneRuler™ DNA Ladder Mix                  | MBI Fermentas  |
| dNTP-Mix                                   | MBI Fermentas  |
| Roti®-Mark STANDARD                        | Roth           |
| Roti®-Mark 10-150 PLUS                     | Roth           |

## 2.6. *Bakterienstämme*

|  |            |
|--|------------|
| <i>Escherichia coli</i> DH5α           | Invitrogen |
| <i>Escherichia coli</i> BL21(DE3)pLysS | Novagen    |

### 3. Methoden

#### 3.1. *Klonierung von RAS- und HST-PCR-Amplifikaten aus Pflanzenmaterial*

##### 3.1.1. Isolierung von genomischer DNA aus Pflanzenmaterial

Unter genomischer DNA (gDNA) versteht man die gesamte (haploide) DNA des Genoms einer Zelle. Das Genom umfasst auch Plastiden- und mitochondriale DNA. Die bei eukaryotischen Genen auftretenden Introns und Exons werden ebenfalls mit dem Begriff Genom erfasst. Introns werden als nicht kodierend bezeichnet, können aber regulatorische Funktionen haben. Introns werden aus der prä-mRNA herausgespleißt. Die resultierende mRNA enthält nur noch kodierende Sequenzen. Durch Variation beim Spleißen kann ein Gen für mehrere Proteinvarianten kodieren. Die unterschiedlichen Sequenzen von Introns und deren Lokalisation können Rückschlüsse auf evolutionäre Veränderungen einzelner Gene zulassen. Bisher sind jedoch nur wenige Funktionen von Introns bekannt. Eine bedeutende Funktion von Introns ist z.B. die Regulation von genomischen Funktionen durch non-coding RNA (Rearick et al., 2011)

Die Isolierung von gDNA aus Blattmaterial erfolgt nach Rogers und Bendich (1985). Das tiefgefrorene Pflanzenmaterial wird in einem mit flüssigem Stickstoff vorgekühlten Mörser zu einem feinen Pulver zermahlen. Danach werden 50 mg des gemörserten Blattmaterials in ein 1,5 ml Eppendorfgefäß überführt, in dem sich 300 µl 2x CTAB-Puffer befinden. Die Herstellung von CTAB-Puffer erfolgte über folgende Rezeptur:

CTAB-Puffer 2x konzentriert:

|          |                                      |
|----------|--------------------------------------|
| 2% (m/V) | Cetyltrimethylammoniumbromid (CTAB)  |
| 100 mM   | Tris/HCl, pH 8,0                     |
| 20 mM    | EDTA                                 |
| 1,4 M    | NaCl                                 |
| 1% (m/V) | Polyvinylpyrrolidon (PVP) (MW 40000) |

Nach Mischen durch leichtes Schütteln wird die Suspension 10 min bei 65 °C inkubiert und anschließend auf Eis abgekühlt. Zur Entfernung von Proteinen und Polysacchariden werden 300 µl Chloroform zugesetzt und gut durchmischt. Durch Chloroform werden Proteine und Polysaccharide ausgefällt und können anschließend abzentrifugiert (5 min, 16000 g) werden. Die obere wässrige Phase mit der gDNA wird nun in ein frisches 1,5 ml Eppendorfgefäß überführt und die gDNA durch Zugabe von 300 µl Isopropanol gefällt. Durch erneute Zentrifugation (5 min, 16000 g) kann die ausgefallene gDNA von dem Überstand getrennt werden. Das Sediment wird mit 500 µl Ethanol (80% v/v) gewaschen und zentrifugiert (5 min, 16000 g). Der Waschschrift wird einmal wiederholt. Die Ethanolüberstände werden vorsichtig mit einer Pipette entfernt. Das gDNA-Sediment wird an Luft getrocknet und in 100 µl destilliertem Wasser gelöst. Die gDNA wird mittels Gelelektrophorese visualisiert.

### **3.1.2. Isolierung von RNA aus Pflanzenmaterial**

RNA entsteht bei der Transkription der DNA (mRNA), wird verwendet als Baustein der Ribosomen (rRNA) und bildet das Gegenstück zum Basentriplett als Transportmolekül für Aminosäuren (tRNA). Bei der Isolierung von RNA ist das Ziel, mRNA zu isolieren und theoretisch das gesamte Transkriptom einer Zelle in Form von bereits gespleißter mRNA zu erhalten. Diese mRNA lässt sich durch reverse Transkriptase in cDNA umschreiben (siehe 3.1.1. cDNA-Synthese) und liefert somit die Grundlage für Amplifikationen von Genen in der PCR.

Zur Durchführung wird die Methode nach Chomczynski und Sacchi (1987) angewendet. Sie ist eine Ein-Schritt-Methode zur Isolierung von Gesamt-RNA aus verschiedenen Zellen. Um zu verhindern, dass die RNA von RNasen abgebaut wird, muss mit Handschuhen gearbeitet werden. Sämtliche Geräte und Lösungen werden doppelt autoklaviert oder zwei Stunden lang bei 200 °C im Trockenschrank „gebacken“. Zur Gewinnung von RNA aus Frischpflanzenmaterial werden die entsprechenden Pflanzenteile geerntet und in flüssigem Stickstoff schockgefroren. In einem mit flüssigem Stickstoff vorgekühlten Porzellanmörser werden die Blätter fein

zermahlen. 50 mg des gemörserten Pflanzenmaterials werden mit 500 µl Lösung D (4 M Guanidinthiocyanat, 25 mM Natriumcitrat, 0,5% Laurylsarcosin; auf pH 7,0 eingestellt) versetzt. Während einer Inkubationszeit von 5 min bei Raumtemperatur wird der Ansatz mehrmals gut gemischt (vortexen). Anschließend werden 50 µl 2 M Natriumacetatlösung (pH 4) zugegeben und durch leichtes Umschwenken gemischt. Danach werden 500 µl Citrat-Puffer-gesättigtes Phenol zugegeben und durch leichtes Umschwenken gemischt. Nach Zugabe von 100 µl eiskaltem Chloroform wird erneut gemischt und die Probe 15 Minuten auf Eis gestellt. Hiernach folgt ein Zentrifugationsschritt (12000 g, 15 min) bei 4 °C. Die obere wässrige Phase wird abgehoben und in ein neues Reaktionsgefäß überführt. 400 µl des Überstands werden mit 400 µl eiskaltem Isopropanol vermischt und 15 min bei -20 °C gelagert, um eine vollständige Fällung der RNA zu erreichen. Die gefällte RNA wird durch Zentrifugation bei 4 °C (12000 g, 15 min) sedimentiert. Der Überstand wird verworfen und Reste über dem Pellet vorsichtig abgesaugt. Das RNA-Pellet wird zunächst mit 70% Ethanol gewaschen. Nach Zentrifugation bei 4 °C (7500 g, 5 min) und Entfernung des Überstandes folgt ein weiterer Waschschrift mit 100% Ethanol. Der Überstand wird nach erneuter Zentrifugation bei 4 °C (7500 g, 5 min) vorsichtig abgehoben und das RNA-Pellet 5 min bei 37 °C getrocknet. Die RNA wird in 20 µl sterilem Wasser bei 50 °C gelöst. Die Integrität der RNA sowie Verunreinigungen mit DNA werden mittels Gelelektrophorese kontrolliert.

### **3.1.3. cDNA-Synthese**

Das virale Enzym „reverse Transkriptase“ synthetisiert an einer mRNA-Vorlage DNA, welche fortan als cDNA bezeichnet wird. Der Vorteil hierbei ist, dass nur die Exons des Genoms amplifiziert werden, da die Introns aus der mRNA schon ausgespleißt sind. mRNA besitzt einen PolyA-Schwanz, der zur Anlagerung der benötigten Primer genutzt wird. Die Oligo-dT-Primer lagern sich an den PolyA-Schwanz an, und die mRNA dient als Vorlage für die DNA-Synthese durch die reverse Transkriptase. Als Produkt entsteht ein cDNA-Strang, der mit der mRNA hybridisiert ist. Diese cDNA kann direkt als Template für die PCR eingesetzt werden. Wichtig beim Arbeiten mit RNA ist die unbedingte RNase-Freiheit der Gefäße und das Arbeiten mit Handschuhen. Schon geringste Mengen RNase zersetzen die RNA, und die Proben

werden unbrauchbar. Die Proben sollten immer auf Eis bearbeitet und mit einem RNase-Inhibitor versetzt werden.

Zur Durchführung wird der RevertAid™ First Strand cDNA Synthesis Kit (Fermentas) verwendet (siehe 2. Material). Dabei werden 0,1-5 µg Gesamt-RNA mit 1 µl Oligo-dT-Primern versetzt und mit doppelt destilliertem Wasser auf 12 µl aufgefüllt. Dieser Ansatz wird bei ca. 4 °C auf Eis gemischt. Anschließend wird 5 Minuten auf 70 °C erhitzt und danach sofort wieder auf Eis gekühlt. Der gekühlte Ansatz wird kurz auf höchster Stufe abzentrifugiert und die folgenden drei Reagenzien hinzupipettiert: 4 µl 5x Reaktionspuffer, 1 µl RiboLock™ RNase Inhibitor (20 U/µl), 2 µl 10 mM dNTP-Mix.

Danach wird der Ansatz vorsichtig durchmischt, abzentrifugiert und 5 min bei 37 °C inkubiert. Nach der Inkubation wird 1 µl Revert Aid™ M-MuLV Reverse Transkriptase (200 U/µl) zugegeben und erneut eine Stunde bei 42 °C inkubiert. Um die Reaktion abzustoppen, wird der Ansatz für zehn Minuten auf 70 °C erhitzt.

#### **3.1.4. Polymerase Kettenreaktion (PCR) zur Amplifizierung von cDNA**

Bei der PCR wird ein DNA-Bereich zwischen zwei bekannten Primer-Sequenzen unter Einsatz thermostabiler Polymerasen und Wechseln der Temperaturen exponentiell vervielfältigt. Der PCR-Ansatz enthält eine thermostabile DNA-Polymerase, die vier dNTPs (dATP, dCTP, dGTP, dTTP), zwei Primer die sich jeweils an die gegenüberliegenden Enden der gesuchten Sequenzen anlagern und die cDNA (Template). Der Ansatz wird zuerst auf 96 °C erhitzt, um das Template einzelsträngig vorliegen zu haben. Danach folgt die Anlagerung der Primer (Annealing) durch schnelles Kühlen auf ca. 50-58 °C (die Temperatur ist abhängig von den eingesetzten Primern). Die komplementären Primer lagern sich an die passende Stelle der einzelsträngigen DNA an. Die Temperaturerhöhung auf ca. 70-72 °C (die Temperatur ist abhängig von der eingesetzten Polymerase) bringt die thermostabile Polymerase auf ihr Temperaturoptimum. Sie kann nun die an der DNA haftenden Primer verlängern, indem sie dNTPs komplementär zum Template

einbaut. Nach der Synthese des neuen DNA Strangs beginnt der Zyklus wieder von vorn. Der Zyklus kann bis zu 40-mal wiederholt werden.

Zur Durchführung wird das GoTaq® Flexi DNA Polymerase Kit von Promega verwendet (siehe 2. Material). Der PCR-Ansatz besteht aus folgenden Reagenzien:

- 5 µl            10 x Colorless GoTaq® Flexi-Puffer
- 3 µl            25 mM MgCl<sub>2</sub>
- 1 µl            10 mM dNTP-Mix
- ca. 1 µl       cDNA
- 0,5 µl        100 pmol/µl Primer forward
- 0,5 µl        100 pmol/µl Primer reverse
- 0,1 µl        (5 U/µl) GoTaq®-Polymerase
- bis auf 25 µl doppelt destilliertes Wasser

Je nach gewünschtem Ergebnis können Temperaturen und Konzentrationen geändert werden. Die Tabellen für PCR-Ansätze gelten hierbei nur als Orientierung. Die PCR wurde zu Beginn der praktischen Arbeit mit dem Stratagene Robocycler oder einem PCR-Cycler der Fa. Eppendorf, später mit einem Biorad-Cycler mit geheiztem Deckel durchgeführt.

Das benutzte Standardprogramm für diese PCR ist in folgende Schritte unterteilt:

| Temperatur              | Zeit    | Zyklen |
|-------------------------|---------|--------|
| 94 °C                   | 2 min   |        |
| 40 °C/42 °C/46 °C/48 °C | 1 min   | 1      |
| 70 °C                   | 1,5 min |        |
| <hr/>                   |         |        |
| 94 °C                   | 0,5 min |        |
| 40 °C/42 °C/46 °C/48 °C | 1 min   | 38     |
| 70 °C                   | 1,5 min |        |
| <hr/>                   |         |        |
| 94 °C                   | 2 min   |        |
| 40 °C/42 °C/46 °C/48 °C | 1 min   | 1      |
| 70 °C                   | 10 min  |        |

Im Anschluss werden die Proben auf 4 °C abgekühlt. Der Erfolg der PCR wird direkt mittels Agarose-Gelelektrophorese untersucht.

### **3.1.5. Isolierung von 5´- und 3´-Enden der kodierenden Gensequenzen eines Organismus – RACE-PCR**

Die RACE-PCR ermöglicht die Isolierung von 5´- und 3´-Enden der mRNA. Dazu wird die gesamte RNA einer Zelle verwendet. Da für die RACE-PCR nur mRNA verwendet werden soll, wird das 5´-Ende aller RNAs außer der mRNA dephosphoryliert. Die mRNA ist durch ihre 5´-Cap-Struktur geschützt. Danach wird die 5´-Cap-Struktur der mRNA entfernt, wobei ein für die Ligation essentieller Phosphatrest stehenbleibt. Dies wird ausgenutzt, um ein RNA-Oligonukleotid mit bekannter Sequenz durch T4-RNA-Ligase nur an die mRNA zu binden. An den Poly-A-Schwanz der mRNA wird über den Oligo-dT-Primer bei der reversen Transkription ebenfalls eine bekannte DNA-Sequenz gebunden. Die nun erhaltene RACE-fertige cDNA besitzt am 5´- und am 3´-Ende bekannte Sequenzen, die als Primerbinderegionen bei der Amplifikation der cDNA-Enden dienen. Als jeweiliger komplementärer Primer dient ein Sequenzfragment eines definierten Proteins, dessen Nukleotidsequenz bereits bekannt ist.

Zur Durchführung wird das GeneRacer™-Kit (Invitrogen) verwendet (siehe 2. Material). Es werden 1-5 µg RNA der zu untersuchenden Pflanze benötigt. Die RNA wird mit Calf Intestinal Phosphatase (CIP) behandelt. Die CIP entfernt 5´-Phosphatreste von Nukleinsäuren. Intakte mRNA mit 5´-Cap-Struktur wird nicht von CIP dephosphoryliert. Dazu wird folgender Ansatz in einem sterilen 1,5 ml Eppendorf Reaktionsgefäß auf Eis gemischt und kurz zentrifugiert:

- X µl            Gesamt-RNA (1-5 µg)
- 1 µl            10x CIP-Puffer
- 1 µl            RNaseOut™ (40 U/µl)
- 1 µl            CIP (10 U/µl)
- 10 µl          DEPC-behandeltes Wasser



Dieser Ansatz wird für 1 h bei 50 °C inkubiert und erneut kurz zentrifugiert. Der Ansatz wird nun auf Eis gestellt. Zur Fällung der RNA werden folgende Reagenzien hinzugefügt und 30 sec gemischt:

- 90 µl DEPC-behandeltes Wasser
- 100 µl Phenol-Chloroform

Der Ansatz wird bei Raumtemperatur zentrifugiert (16000 g, 5 min), und die obere wässrige Phase wird in ein neues Eppendorf Reaktionsgefäß überführt. Zu diesem Überstand werden folgende Reagenzien hinzugefügt und durch Umschwenken gut gemischt:

- 2 µl Muschelglykogen (10 mg/ml)
- 10 µl 3 M Natriumacetat pH 5,2
- nach Mischen 220 µl 95% Ethanol

Der Ansatz wird geschüttelt und 10 min auf Eis inkubiert, um die RNA auszufällen. Die ausgefallene RNA wird bei 4 °C abzentrifugiert (16000 g, 20 min) und der Überstand vorsichtig mit einer Pipette entfernt. Die RNA wird mit 500 µl 70% Ethanol gewaschen. Nach mehrmaligem vorsichtigem Mischen durch Umschwenken wird erneut bei 4 °C zentrifugiert (16000 g, 2 min). Das Ethanol wird vorsichtig mit einer Pipette entfernt und das RNA-Pellet bei Raumtemperatur 2 min getrocknet. Die RNA wird in 7 µl DEPC-behandeltem Wasser gelöst (es können auch 8 µl DEPC-behandeltes Wasser verwendet werden, um 1 µl Probe zur Kontrolle in einer Gelelektrophorese zu erhalten.)

Die nun dephosphorylierte RNA und die intakte mRNA mit 5'-Cap-Struktur werden nun mit Tobacco Acid Phosphatase (TAP) behandelt. TAP entfernt die 5'-Cap-Struktur der intakten mRNA und hinterlässt 5'-Phosphatenden, die für die Ligation des RNA-Oligonukleotids verwendet werden. Die restliche, im ersten Schritt dephosphorylierte RNA wird nicht mit dem Oligonukleotid verknüpft. Dazu wird auf Eis folgender Ansatz in einem sterilen 1,5 ml Eppendorf Reaktionsgefäß verwendet:

- 7 µl dephosphorylierte RNA/ intakte mRNA
- 1 µl 10x TAP-Puffer
- 1 µl RNaseOut™ (40 U/µl)
- 1 µl TAP (0,5 U/µl)

Dieser Ansatz wird gemischt, abzentrifugiert und 1 h bei 37 °C inkubiert. Nach der Inkubation wird wieder abzentrifugiert, die Reaktion auf Eis gesetzt und die RNA erneut gefällt (s.o.).

An die 5'-Phosphatenden der intakten mRNA wird nun das RNA-Oligonukleotid (GeneRacer™-RNA-Oligo) geknüpft. Dazu werden zum lyophilisierten RNA-Oligonukleotid (0,25 µg) mit 44 Basen

5'-CGACUGGAGCAGCACGAGGACACUGACAUGGACUGAAGGAGUAGAAA-3'

7 µl der gefällten RNA zugegeben. Der Ansatz wird durch Pipettieren gut gemischt, abzentrifugiert und bei 65 °C 5 min inkubiert. Danach wird für 2 min auf Eis gekühlt und die folgenden Reagenzien mit der RNA vermischt:

- 1 µl 10x Ligase Puffer
- 1 µl 10 mM ATP
- 1 µl RNaseOut™ (40 U/µl)
- 1 µl T4 RNA Ligase (5 U/µl)

Der Ansatz wird 1 h bei 37 °C inkubiert, abzentrifugiert und die RNA erneut gefällt (s.o.).

Nun erfolgt die reverse Transkription, nach der man RACE-fertige cDNA mit bekannten Primerbereichen an den 5'- und 3'-Enden erhält. Der hierfür eingesetzte Oligo-dT-Primer (GeneRacer™-Oligo-dT Primer) mit 60 Basen

5'-GCTGTCAACGATACGCTACGTAACGGCATGACAGTG (T)<sub>24</sub>-3'

bindet an den Poly-A-Schwanz der mRNA und ist an seinem 3'-Ende um eine bekannte Sequenz von 36 Basen verlängert. Dazu werden die folgenden Reagenzien mit der RNA gemischt:

- 1 µl GeneRacer™ Oligo-dT-Primer
- 1 µl dNTP Mix (10 mM)
- 1 µl steriles destilliertes Wasser

Der Ansatz wird für 5 min auf 65 °C erhitzt, 1 min auf Eis gekühlt und abzentrifugiert. Danach werden folgende Reagenzien hinzugefügt:

- 4 µl 5x First Strand Puffer
- 1 µl 0,1 M DTT
- 1 µl RNaseOut™ (40 U/µl)
- 1 µl SuperScript™ III Reverse Transkriptase (200 U/µl)

Der Ansatz wird durch Pipettieren gemischt, abzentrifugiert und 60 min bei 50 °C inkubiert. Die Reverse Transkriptase wird durch Erhitzen auf 70 °C für 5 min inaktiviert. Nach Abkühlen auf Eis für 2 min wird der Ansatz abzentrifugiert.

Zum Verdauen der verbliebenen RNA wird 1 µl RNase H (2 U) zugefügt und 20 min bei 37 °C inkubiert. Die cDNA ist nun einsatzbereit für die RACE-PCR, bei der mit genspezifischen und RACE-Primern die gewünschten 3'- bzw. 5'-Enden einer cDNA vervielfältigt werden können.

### 3.1.6. Agarose-Gelelektrophorese

Mit Hilfe der Agarose-Gelelektrophorese können DNA- und RNA-Stränge ihrer Größe nach aufgetrennt werden. Durch das angelegte elektrische Feld wandern die Nukleinsäuren aufgrund ihrer negativen Nettoladung zur Anode. Dabei wandern größere Moleküle langsamer durch das Gel, weil sie durch das ausgebildete Agarosenetz stärker zurückgehalten werden. Kleinere Moleküle können besser mit höheren Agarosekonzentrationen getrennt werden, wobei Agarosekonzentrationen von 0,5-2,0 % verwendet werden.

Zur Durchführung wird eine entsprechende Menge Agarose (0,5-2,0% w/v) in Elektrophoresepuffer, versetzt mit einer 1%igen Ethidiumbromid-Stammlösung (Phenanthridinium-3,8-diamino-5-ethyl-6-phenylbromid) (Abb. 3.1) durch Erhitzen

gelöst und in einen Elektrophoreseschlitten gegossen. In das Gel wird ein Kamm gesteckt, der die Probestaschen in das härtende Gel einbringt. Das Gel muss bis zum völligen Erstarren abkühlen. Danach wird das Gel in die Elektrophoresekammer gesetzt und mit Elektrophoresepuffer überschichtet. Als Elektrophoresepuffer wird 0,5x TBE-Puffer (45 mM Tris, 45 mM Borsäure, 1,5 mM Na-EDTA) verwendet. Nach dem Überschichten mit Elektrophoresepuffer wird der Kamm aus dem Gel gezogen, und die Proben können in die Taschen gefüllt werden. Die Proben müssen zuvor mit Probenpuffer versetzt werden, damit deren spezifische Dichte erhöht wird und sie dadurch in die Probestaschen einsinken. Zur Erhöhung der Probendichte wird Glycerin verwendet. Als Probenpuffer dient 6x Loading Dye (10 mM Tris/HCl (pH 7,6), 60 mM EDTA, 60% Glycerin, je 0,03% Bromphenolblau und Xylencyanol FF). Zusätzlich zu den Proben wird ein Größenmarker aufgetragen, um spätere Banden einem definierten Größenbereich zuordnen zu können. Der verwendete Größenmarker ist der GeneRuler™ DNA Ladder Mix von Fermentas. Die DNA bzw. RNA wird unter UV-Licht (254 nm) durch das angelagerte Ethidiumbromid visualisiert und das Ergebnis mittels Fotografie dokumentiert. Ethidiumbromid ist ein Fluoreszenzindikator, der in DNA und RNA interkaliert. Dabei können kovalente Bindungen zwischen DNA und Ethidiumbromid entstehen, weshalb ein Kontakt mit der Substanz zwingend vermieden werden muss. Bei Anregung durch UV-Licht leuchten Stellen mit Nukleinsäuren orange auf.

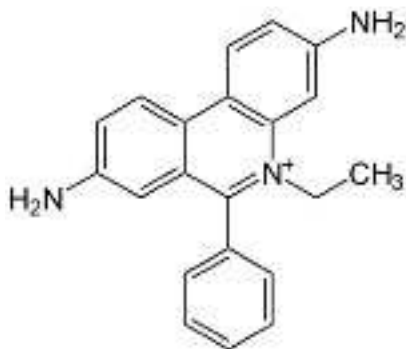


Abb. 3.1: Strukturformel von Ethidiumbromid (Quelle: <http://de.wikipedia.org/wiki/Ethidiumbromid>)

Um die DNA der Banden anschließend isolieren zu können, werden die Banden mit einem Skalpell ausgeschnitten und die Agaroseblöckchen in ein Eppendorfgefäß überführt.

### 3.1.7. DNA-Extraktion aus Agarosegelen

Durch die Bindung von DNA an eine Silicamembran, die in einem Aufsatz für Eppendorf Reaktionsgefäße steckt, kann nach Auflösen der Agarose die DNA in ausreichender Reinheit gewonnen werden. Dabei wird durch chaotrope Salze und Isopropanol die DNA aus der wässrigen Phase gedrängt und bindet mit dem negativ geladenen Zuckerphosphatrückgrat an die positiv geladenen Silanoloberflächen. Der Reinheitsgrad der erhaltenen DNA ermöglicht die direkte Weiterbearbeitung in enzymatischen Reaktionen (z.B. Ligationen).

Zur Durchführung wird der NucleoSpin<sup>®</sup> Extract II Kit (Macherey-Nagel) verwendet. Die nach der erwarteten Größe identifizierte DNA-Bande wird unter UV-Licht aus dem Agarosegel herausgeschnitten. Es ist darauf zu achten, dass keine Fremdbanden mit ausgeschnitten werden und die Dauer der UV-Exposition möglichst kurz ist. Das ausgeschnittene Gelstück wird in ein 1,5 ml Eppendorf Reaktionsgefäß überführt und mit ca. der doppelten Menge an NT-Puffer versetzt. Durch Erwärmen auf 50 °C für 10 min und gelegentliches Schütteln wird die Agarosegelmatrix aufgelöst. Nach der Inkubation, wenn sich im NT-Puffer keine Schlieren mehr zeigen, wird die Lösung auf ein NucleoSpin-Säulchen gegeben und zentrifugiert (11000 g, 1 min). Die DNA bindet an die Säule und wird weiter gereinigt, indem in das Säulchen 600 µl NT3-Puffer gegeben und erneut zentrifugiert wird (11000 g, 1 min). Der NT3-Puffer muss vollständig von der Säule entfernt werden. Dafür wird die Säule erneut zentrifugiert (11000 g, 2 min). Um die DNA von der nun trockenen Säule zu eluieren, wird doppelt destilliertes Wasser direkt auf die Silicaoberfläche pipettiert und eine Minute inkubiert. Durch Zentrifugation wird die DNA eluiert (11000 g, 1 min). Das Elutionsvolumen (15-50 µl Wasser oder Elutionspuffer) richtet sich nach der Fragmentgröße und der gewünschten Aufkonzentrierung. Je kürzer die aufzureinigende DNA ist, desto weniger Wasser wird zur quantitativen Elution benötigt. Das DNA-Eluat kann bis zur weiteren Verwendung bei -20 °C gelagert werden.

### 3.1.8. Ligation

Unter einer Ligation versteht man die enzymkatalysierte Verknüpfung zweier DNA-Fragmente, so wird z.B. ein DNA-Fragment in ein Plasmid eingebunden. Plasmide sind zirkuläre, doppelsträngige DNA-Moleküle, die sich unabhängig vom bakteriellen Genom in Bakterien vermehren können. Sie bestehen mindestens aus einem Replikationsstart (origin of replication, ori) einem Selektionsgen (hier Ampicillin-Resistenzgen, ampR) und einer Klonierungsstelle, der sogenannten multiple cloning site (mcs).

Zur Durchführung wurden verschiedene Vektorsysteme verwendet (siehe 2. Material). Im Allgemeinen setzt sich der Ligationsansatz aus folgenden Bestandteilen zusammen:

- 5 µl 2x Ligations-Puffer
- 1 µl Vektor
- 3 µl gereinigtes PCR-Produkt
- 1 µl T4 DNA Ligase (3 U/µl)

Ligase sollte nicht im Überschuss eingesetzt werden, da sich das Enzym autoinhibiert. Um eine effiziente Ligation zu erzielen, wird der Ligationsansatz über Nacht bei 4 °C inkubiert.

### 3.1.9. Herstellung kompetenter Zellen

Kompetenz ist die Fähigkeit von prokaryotischen Zellen, DNA aufzunehmen, die frei im umgebenden Medium vorliegt. Eine Möglichkeit, Bakterien kompetent zu machen, ist, sie bestimmten Chemikalien auszusetzen. Dabei ist eine Standardmethode, die Kompetenz unter Verwendung einer hochkonzentrierten Calciumchloridlösung herbeizuführen.

Zur Durchführung wird der Roti<sup>®</sup>-Transform Kit (Carl Roth) verwendet (siehe 2. Material). Unter sterilen, antibiotikafreien Bedingungen wird eine Übernachtskultur des gewünschten *E. coli*-Stamms in LB-Medium angelegt. Das LB-Medium setzt sich wie folgt zusammen:

- 10 g Trypton
- 5 g Hefeextrakt
- 10 g Natriumchlorid
- auffüllen mit destilliertem Wasser auf 1 l
- bei Festmedien werden 15 g/l Agar-Agar zugesetzt

Das fertig gemischte Medium wird auf pH 7 eingestellt und autoklaviert.

Hier wird der *E. coli* Stamm DH5 $\alpha$  verwendet (Genotyp: *huA2*  $\Delta$ (*argF-lacZ*)U169 *phoA glnV44*  $\Phi$ 80  $\Delta$ (*lacZ*)M15 *gyrA96 recA1 relA1 endA1 thi-1 hsdR17*). Die Übernachtskultur wird verwendet, um 50 ml LB-Medium mit ca. 500  $\mu$ l anzuimpfen. Die Bakterien werden nun weiter bei 37 °C im Schüttler (220 UpM) inkubiert, bis sie eine OD<sub>600</sub> zwischen 0,4 und 0,7 erreicht haben. Ist die geforderte Bakteriendichte erreicht, werden zehnmal je 1,9 ml der Bakteriensuspension in sterile 2 ml Eppendorfgefäße überführt und zentrifugiert (3000 g, 5 min). Die sedimentierten Bakterien werden in 100  $\mu$ l eiskaltem LB-Medium resuspendiert. Danach wird 90  $\mu$ l eiskaltes Roti<sup>®</sup>-Transform 1 zugegeben und vorsichtig gemischt. Als letzter Zusatz wird 10  $\mu$ l Roti<sup>®</sup>-Transform 2 hinzupipettiert und erneut vorsichtig gemischt. Die Ansätze werden 5 Minuten auf Eis inkubiert. Die nun kompetenten Zellen können direkt für Transformationen verwendet oder für spätere Transformationen bei -80 °C gelagert werden.

### **3.1.10. Transformation**

Unter Transformation versteht man die passive Aufnahme von freier DNA in Bakterienzellen. Die kompetenten Zellen besitzen eine perforierte Membran, durch die Plasmide diffundieren können. Die Plasmide liegen neben dem Bakterienchromosom autonom im Cytosol der Bakterienzellen vor. Wegen der

perforierten Membran sind ein vorsichtiges Pipettieren und ein vorsichtiges Mischen der Bakteriensuspensionen notwendig.

Zur Durchführung werden 5 µl des Ligationsansatzes mit kompetenten Bakterien vermischt und für 30 bis 60 min auf Eis inkubiert. Im Anschluss werden die transformierten Bakterien auf Ampicillin-haltigen LB-Agar-Platten ausplattiert und über Nacht bei 37 °C inkubiert. Nach Aufkochen von LB-Festmedium und Abkühlen auf unter 60 °C kann Ampicillin (100 µg/ ml) direkt zum verflüssigten Medium zugegeben oder bei schon fertig gegossenen Platten durch Ausspateln zugefügt werden. Durch den Selektionsmarker Ampicillin (Abb. 3.2) werden nur von denjenigen Bakterien Kolonien gebildet, die mit der Aufnahme des Plasmids die Ampicillinresistenz erworben haben.

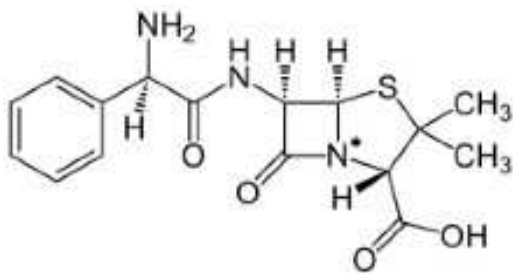


Abb. 3.2: Strukturformel des  $\beta$ -Lactam-Antibiotikums Ampicillin mit Markierung des reaktiven Amins im gespannten Ring (\*) (Quelle: <http://en.wikipedia.org/wiki/Ampicillin>, modifiziert).

Ampicillin zählt zu den Breitbandantibiotika und greift, wie alle  $\beta$ -Lactam-Antibiotika, in die bakterielle Zellwandsynthese ein. Das Ampicillin-Resistenzgen *ampR* stammt aus dem Bakterium *Salmonella paratyphi* und sorgt für die Bildung des Enzyms  $\beta$ -Lactamase, durch dessen Wirkung Ampicillin inaktiviert wird.

Ferner wurden pro Platte 100 µl 0,1 M Isopropyl- $\beta$ -D-thiogalactopyranosid (IPTG)-Lösung, welches als künstlicher Induktor des Lactose-Operons in *E. coli* verwendet wird, und 10 µl X-Gal (5-Brom-4-chlor-3-indoxyl- $\beta$ -D-galactopyranosid) ausgestrichen.



Das Enzym  $\beta$ -Galactosidase hydrolysiert X-Gal zu Galactose und 5-Brom-4-chlor-indoxyl (5-Brom-4-chlor-3-hydroxyindol). Das 5-Brom-4-chlor-indoxyl wird vom Sauerstoff der Luft zum tiefblauen Farbstoff 5,5'-Dibrom-4,4'-dichlor-indigo oxidiert.

Somit werden die Klone, die nicht transformiert wurden, blau; die Klone, bei denen die Transformation funktioniert hat, bleiben weiß.

### **3.1.11.      Aufreinigung von Plasmid-DNA**

Für die Präparation von Plasmiden werden die Bakterien durch alkalische Lyse aufgeschlossen. Zur Aufreinigung der Plasmid-DNA werden die meisten anderen zellulären Moleküle zuerst ausgefällt und abzentrifugiert. Der dafür verwendete Puffer enthält zusätzlich RNase, um störende RNA zu zersetzen. Nach der Zentrifugation wird der Überstand auf eine Silicasäule geladen (siehe 3.1.5 DNA Extraktion aus Agarosegelen). Die gebundene Plasmid-DNA wird gewaschen und eluiert. Danach kann die Plasmidkonzentration photometrisch bestimmt und die Plasmid-DNA bei -20 °C gelagert werden.

Zur Durchführung wird der QIAprep® Spin Miniprep Kit (Qiagen) verwendet (siehe 2. Material). Die Plasmid-DNA wird aus 3 ml einer Übernachtskultur von transformierten *E.coli*-Klonen gewonnen. 1,5 ml der Kultur werden in ein 1,5 ml Eppendorf Reaktionsgefäß überführt und zentrifugiert (3000 g, 5 min). Der Überstand wird verworfen, und zum Pellet werden weitere 1,5 ml Bakteriensuspension gegeben und die Zentrifugation (3000 g, 5 min) wiederholt. Das Pellet wird in 250  $\mu$ l P1 Puffer resuspendiert und 250  $\mu$ l P2 Puffer zugegeben. Die Gefäße werden anschließend zehnmal über Kopf gedreht, um die Suspension zu durchmischen. Diese vorsichtige Vorgehensweise ist notwendig, da die Plasmid-DNA in dem nun stark alkalischen Milieu sehr scherempfindlich ist. Nach höchstens 5 min Inkubation wird mit der Zugabe von 350  $\mu$ l N3 Puffer die Suspension neutralisiert. Nach erneutem Durchmischen erfolgt die Präzipitation der unerwünschten Bestandteile. Dabei bildet sich ein weißer Niederschlag in der Suspension. Das Präzipitat wird abzentrifugiert (16000 g, 10 min). Der Überstand wird auf die Säule mit der eingesetzten Silicamembran überführt und zentrifugiert (16000 g, 1 min). Die am Säulenmaterial adsorbierte Plasmid-DNA wird mit 500  $\mu$ l PB-Puffer gewaschen und erneut

zentrifugiert (16000 g, 1 min). Es folgt ein zweiter Waschschriff mit 750 µl PE-Puffer und Zentrifugation (16000 g, 1 min). Der PE-Puffer muss vollständig von der Säule entfernt werden. Dafür wird die Säule erneut zentrifugiert (16000 g, 2 min). Um die DNA von der nun trockenen Säule zu eluieren, wird 50 µl doppelt destilliertes Wasser direkt auf die Silicaoberfläche pipettiert und eine Minute inkubiert. Danach wird die Säule in ein steriles 1,5 ml Eppendorf Reaktionsgefäß gesetzt, und durch Zentrifugation wird die DNA eluiert (16000 g, 1 min).

### 3.1.12. Verdau mit Restriktionsenzymen

Restriktionsendonukleasen sind Enzyme, die DNA an festgelegten Positionen (Restriktionsschnittstellen) schneiden. Sie wurden zuerst in Bakterien entdeckt, die diese nutzen, um sich vor Fremd-DNA zu schützen. Man unterscheidet drei Grundtypen, die Sequenzen an unterschiedlichen, teils willkürlichen Stellen schneiden. Wichtig für die Arbeiten im Labor ist der Typ II, der die Sequenz nahe oder innerhalb der Erkennungssequenz schneidet. Die meisten Restriktionsendonukleasen erkennen dabei palindromische Sequenzen, deren beide DNA-Stränge die gleiche Sequenz, von 5' nach 3' gelesen, aufweisen. Als Beispiel dient die in dieser Arbeit häufig verwendete Restriktionsendonuklease *EcoRI*, die die folgende Sequenz erkennt (Abb. 3.3).

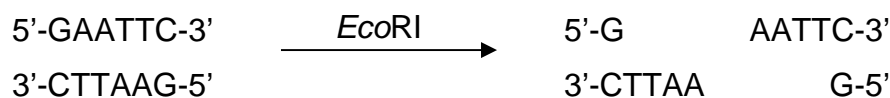


Abb. 3.3: Schematische Darstellung der Aktivität der Restriktionsendonuklease *EcoRI*

Die Erkennungssequenzen können mittels PCR an DNA-Fragmente angehängt und danach zur Ligation in Vektoren genutzt werden. Interne Schnittstellen, die auf Restriktionskarten dargestellt werden, können zur indirekten Überprüfung der Sequenz verwendet werden. Wird die Sequenz an der/den spezifischen Stelle/n geschnitten und es entstehen zwei oder mehrere Fragmente, die den zuvor

theoretisch bestimmten Fragmentgrößen gleichen, ist die geschnittene Sequenz mit hoher Wahrscheinlichkeit die gesuchte.

Zur Durchführung wurden verschiedene Restriktionsendonukleasen verwendet (siehe 2. Material). Im Allgemeinen setzt sich der Verdau aus folgenden Bestandteilen zusammen:

- 5 µl aufgereinigtes Plasmid
- 3 µl für die eingesetzten Restriktionsendonukleasen spezifischer 10x Puffer
- 0,5 µl Restriktionsendonuklease 1 (10 U/µl)
- 0,5 µl Restriktionsendonuklease 2 (10 U/µl)
- 6-6,5 µl destilliertes Wasser

Dieser Ansatz wird für 3 Stunden bei 37 °C inkubiert, bevor er über ein Agarosegel analysiert wird.

### **3.1.13. Photometrische DNA-Quantifizierung**

Die DNA-Konzentration wird photometrisch durch Absorptionsmessung bei 260 nm bestimmt. Aus dem Verhältnis der Absorption von 260 zu 280 nm wird ein Quotient gebildet, der einen Hinweis auf die Reinheit und potentielle Verunreinigung mit Protein gibt. Eine ausreichend proteinfreie Nukleinsäurelösung weist ein Verhältnis  $OD_{260}/OD_{280}$  von 1,8 bis 2,0 auf. Diese Konzentrationsbestimmung ist als Richtwert zu sehen und sollte mittels Abschätzung der DNA-Mengen im Agarosegel verifiziert werden. Eine kontinuierliche Kalibrierung der photometrischen Ergebnisse über Konzentrationsstandards ist dabei von Vorteil

Zur Durchführung wird die Plasmid-DNA in autoklaviertem Wasser verdünnt und gegen Wasser als Nullwert gemessen. Liegt die Konzentration außerhalb des Messbereichs des Photometers muss die Plasmid-DNA erneut verdünnt werden.

#### **3.1.14. DNA-Sequenzierung**

Die Nukleotidsequenz der aus den PCR-Reaktionen gewonnenen Fragmente wird über Sequenzierung endgültig bestimmt. Die Sequenzierung wird extern vorgenommen von der Firma Eurofins MWG Operon. Dazu werden gereinigte Plasmid-DNA (50 bis 100 ng/µl) in einem Gesamtvolumen von >15 µl eingeschickt.

Das Sequenzierungsergebnis wird einer Datenbankrecherche (<http://blast.NCBI.nlm.nih.gov/Blast.cgi>) unterzogen. Dabei wird die Sequenz mit schon bekannten Sequenzen verglichen und Proteine und Domänen mit hohen prozentualen Übereinstimmungen angegeben.

#### **3.1.15. Kryokonservierung von Prokaryoten**

Die Bakterienkolonien, die den pGEM<sup>®</sup>-T-Vektor mit der gewünschten DNA enthalten, sollen dauerhaft gelagert werden. Dazu eignet sich am besten eine Bakteriendauerkultur.

Dazu werden 425 µl einer frischen Übernachtskultur mit 75 µl sterilem Glycerin vermischt. Glycerin fungiert als Gefrierschutz für die Bakterien. Diese können nach Schockfrieren in flüssigem Stickstoff bei -80 °C über mehrere Jahre gelagert werden. Aus einer Bakteriendauerkultur kann jederzeit wieder eine Übernachtskultur angeimpft werden.

## **3.2. Expressionsversuche der gefundenen RAS und HST cDNAs**

### **3.2.1. Expression der cDNA**

Die amplifizierte und modifizierte vermeintliche Volllängensequenz aus *Symphytum officinale* wird in *E. coli* BL21(DE3)pLysS mittels des Expressionsplasmids pET 15b zur Expression gebracht. Die cDNA ist zusätzlich mit Restriktionsschnittstellen und Aufreinigungstags versehen.

Zur Durchführung wird ein Kolben mit 100 ml LB-Ampicillin-Medium mit 1 ml Übernachtskultur versetzt. Die Übernachtskultur wurde entweder mit frisch transformierten oder bereits kryokonservierten Klonen angeimpft. Dieser Ansatz wird bei 37 °C und Schütteln (200 U/min) bis zu einer optischen Dichte (bei 600 nm) zwischen 0,5 und 0,8 inkubiert. Die Bakteriensuspension wird nun auf zwei Kolben mit 50 ml Volumen aufgeteilt. Zu einem der beiden Kolben wird eine Endkonzentration von 1 mM IPTG zugegeben und sofort geschwenkt. Die zweite Bakterienkultur wird nicht induziert und dient später als Vergleichsprobe. Nach Induktion wird die Bakteriensuspension weiter inkubiert und nach 5 h geerntet. Zur Ernte werden die Kolben sofort auf Eis gestellt und die Bakterien bei 4 °C und 5000 g sedimentiert. Das Arbeiten auf Eis ist notwendig, da aggressive bakterielle Proteasen das exprimierte Protein zerstören können. Der Überstand wird dekantiert und die Reste vorsichtig mit einer Pipette oberhalb des Pellets abgesaugt. Die Pellets werden gewogen und in flüssigem Stickstoff schockgefroren; sie können bis zur weiteren Verarbeitung bei -80 °C gelagert werden. Die gefrorenen Bakteriensedimente werden langsam auf Eis aufgetaut und in der vierfachen Menge der Bakterienmasse KPi-Puffer (0,1 M, pH 7,0) resuspendiert. Durch Ultraschallaufschluss mit 0,3 Zyklen und 100% Leistung für 1 min werden die Bakterien zerstört. Dabei wird darauf geachtet, stetig die Proben mit Eis gekühlt zu halten. Durch Zentrifugation (11000 g, 5 min) werden Zellbruchstücke abgetrennt. Der gewonnene Überstand enthält u.a. das gelöste Protein.

Mittels Gelelektrophorese wird die Anwesenheit des Zielproteins überprüft. Eine sichtbare Bande bei der gesuchten Größe ist ein guter Hinweis auf das Gelingen der Expression. Das Protein kann aber auch sehr schlecht exprimiert und dann über Gelelektrophorese nicht detektierbar sein. Dann können spezifische Enzymaktivitätstests einen Hinweis auf Anwesenheit des Proteins geben.

### 3.2.2. Optimierung der Expression

Die Expression kann durch Variation verschiedener Parameter optimiert werden. Ziel ist eine Überexpression des Proteins unter Beibehaltung der biologischen Aktivität. Die Optimierung der Expressionsleistungen der einzelnen Klone während der Umsetzung dieser Arbeit zeigte jedoch die Limitierung des Systems bei nicht bekannten Proteinen. Die Expressionsbakterien konnten die Proteine trotz zahlreicher Variationen der Kulturbedingungen nie in ausreichenden Mengen, bezogen auf die Sichtbarkeit von Proteinbanden in SDS-Gelen, herstellen.

Zur Optimierung der Expression der Proteine wurden folgende Versuche durchgeführt:

- Variation des Erntezeitpunkts nach Inkubation

Hierbei wird der optimale Erntezeitpunkt ermittelt. Dazu werden mehrere Kolben mit Übernachtskultur angeimpft. Liegt die optische Dichte zwischen 0,7 bis 0,8 wird durch Zugabe von 1 mM IPTG induziert. Eine Kultur wird direkt nach Induktion geerntet. Alle weiteren Kulturen werden bei 37 °C geschüttelt und nacheinander geerntet. Die Erntezeitpunkte liegen zwischen 1 und 24 h nach Induktion.

- Variation der Inkubationstemperatur

Erniedrigung der Temperatur kann zu einer Erhöhung des Anteils an gelöstem Protein und erhöhter biologischer Aktivität führen. Drei Bakterienkulturen werden bis zur gewünschten optischen Dichte bei 37 °C kultiviert. Nach Induktion mit 1 mM IPTG werden die Kulturen bei 37 (5 h), 25 (10 h) und 4 °C (>96 h) für die jeweils angegebene Zeit kultiviert.

- Variation der IPTG-Konzentration

Die Expressionsstärke im pET-System kann über die zugegebene Menge an IPTG beeinflusst werden. Dabei wird die IPTG-Konzentration gesucht, die ein Optimum zwischen exprimiertem und gelöstem, biologisch aktivem Protein erzielt. Laut Herstellerangaben können niedrige Expressionslevel die Löslichkeit der gebildeten Proteine erhöhen. Es wird empfohlen, IPTG-Konzentrationen zwischen 25 µM und 1 mM zu testen.

➤ Variation der Induktion

Auf die Wachstumsphase bei 37 °C wird verzichtet und drei Hauptkulturen sofort bei 25 °C geschüttelt. Die Kulturen werden ohne Induktion kultiviert und nach 17 h, 28 h und 42 h geerntet.

➤ Variation des Aufschlusspuffers

Detergenzien im Aufschlusspuffer können in einigen Fällen den Anteil an gelöstem Protein im Rohextrakt erhöhen. Zur Herstellung des Proteinrohextrakts wird ein Aufschlusspuffer mit Detergenz verwendet:

- 292,2 mg            NaCl
- 37,2 mg            EDTA
- 200 mg            Desoxycholat
- 1 ml                Triton X-100
- Ad 100 ml        50 mM Tris/HCl pH 7,5

Der Aufschlusspuffer der Kontrollprobe ist 0,1 M KPi pH 7,0. Zur Herstellung eines weiteren Rohextraktes wird His-Tag-Bindepuffer verwendet. Würde ein Aufschluss gelingen, wäre ein Umpuffern der Proteinfragmente vor der His-Tag-Aufreinigung nicht mehr notwendig.

➤ Variation der Aufschlusszeit

Die Bakterien werden mehrere Minuten aufgeschlossen. Nach jeder Minute wird eine Probe gezogen und deren Proteingehalt bestimmt. Es wird in einem Diagramm die Aufschlusszeit gegen den Proteingehalt aufgetragen. Erreicht der Proteingehalt ein Plateau, ist davon auszugehen, dass alle Zellen aufgeschlossen sind.

➤ Variation des Expressionszellstamms

Die Kombination aus Vektor und *E. coli*-Stamm beeinflusst die Expressionsstärke des Zielproteins. In dieser Arbeit wurden die Expressionsstämme *E. coli* JM109 und *E. coli* BL21(DE3)pLysS verwendet.

Zu einer verbesserten Überexpression werden verschiedene Übernachtskulturen angeimpft.

### 3.2.3. Herstellung von Proteinrohextrakt

Es konnte ein Expressions- und Aufreinigungsverfahren entwickelt werden, bei dem die Proteine in ausreichender Menge hergestellt werden, um spätere Enzymtests (siehe 3.3. Bestimmung der Aktivität der heterolog exprimierten Enzyme RAS und HST) durchführen zu können.

Folgende Durchführung wurde standardmäßig nach vorherigen Optimierungsversuchen (siehe 3.2.2. Optimierung der Expression) umgesetzt.

Aus Glycerolstocks werden in 5 ml LB-Ampicillin-Medium Übernachtskulturen angesetzt. Mit Übernachtskulturen des Expressionsstamms *E. coli* BL21(DE3)pLysS, der die gewünschte cDNA in dem Vektor pET-15b enthält, werden die Hauptkulturen angeimpft. Bei jeder Hauptkultur wird so verfahren, dass 100 ml LB-Ampicillin-Medium mit 4 ml Übernachtskultur versetzt und die 250 ml-Kolben bei 37 °C und 220 UpM für 1,5 h geschüttelt werden. Die Induktion erfolgt bei einer OD<sub>600</sub> von 0,4-1,0. Dafür wird jeder Ansatz mit 1 ml 0,1 M IPTG induziert. Nach weiteren 5 h bei 37 °C werden die Kulturen geerntet.

Dabei werden die Bakterien bei 4 °C abzentrifugiert (5000 g, 5 min). Das Bakterienpellet wird mit 20 mM Tris/HCl pH 8,0 aufgeschwemmt und erneut zentrifugiert (5000 g, 5 min). Die Bakterienpellets werden weiter, nachdem sie einmal eingefroren wurden, zu Rohextrakten verarbeitet. Jedes Pellet wird gewogen und pro Gramm in 5 ml 0,1 M KPi-Puffer oder His-Tag-Bindepuffer (0,05 M KPi-Puffer, pH 8,0; 300 mM NaCl), resuspendiert. Es folgen Aufschluss und Zentrifugation unter Standardbedingungen (0,3 Zyklen, 100%, 3x 30 sec Puls und 30 sec zum Kühlen auf Eis). Der Ansatz wird bei 4 °C zentrifugiert (10000 g, 10 min) und der gewonnene Überstand auf Eis gestellt. Überstände mit den gleichen rekombinanten Proteinen können nun vereinigt werden.



### 3.2.4. Synthese der CoA-aktivierten Kaffeesäure (Caf-CoA) bzw. 4-Cumarsäure (pC-CoA)

Die Synthese der CoA-aktivierten Säuren beginnt mit der Herstellung der entsprechenden Succinimid-Ester. Diese sind bereits in der Laborsammlung vorhanden, so dass direkt mit der Bildung der Thioester begonnen werden kann.

Dazu wurden zunächst 25,2 mg  $\text{NaHCO}_3$  in ein hohes Reagenzglas gefüllt, in 6 ml destilliertem Wasser gelöst und 15 min mit Stickstoff begast. In diesem basischen Milieu werden 24,6 mg Coenzym A Trilithiumsalz-Dihydrat (CoA) gelöst und weitere 15 min mit Stickstoff begast. Nun werden 56 mg des entsprechenden Succinimid-Esters, gelöst in 0,5 ml Aceton p.a., langsam zur CoA-Lösung getropft. Der sich bildende Niederschlag wird durch Addition weiterer 6 ml Aceton p.a. aufgelöst. Das Reaktionsgefäß mit der nun klaren gelben Lösung wird mit Nescofilm verschlossen, mit Alufolie abgedunkelt und über Nacht bei Raumtemperatur inkubiert. Am folgenden Tag wird das Aceton durch einen über die Lösung geleiteten Stickstoffstrom vollständig entfernt. Danach wird 10 min bei 1000 g zentrifugiert. Der Überstand, der das gewünschte Reaktionsprodukt enthält, wird über Nacht chromatographisch aufgereinigt. Verwendet werden dazu DC-Celluloseplatten mit Fluoreszenzindikator ( $F_{254}$ ) und als Fließmittel eine Mischung aus n-Butanol/Essigsäure/Wasser im Verhältnis 4:2:3. Unter UV-Licht bei 254 nm können die CoA-Thioester durch Fluoreszenzlöschung als dunkle Banden identifiziert und markiert werden. Die Retentionsfaktoren sind 0,42 für 4-Cumaroyl-CoA (pC-CoA) und 0,3 für Caf-CoA. Die Trägerschicht wird im Bereich der Banden ausgekratzt. Zur Elution der CoA-Thioester wird so lange mit Wasser ausgeschüttelt und nach Herunterzentrifugieren die Überstände photometrisch (200 nm bis 500 nm) vermessen bis keine CoA-Ester mehr detektierbar sind. Die Überstände mit den gewünschten Reaktionsprodukten werden vereinigt und Identität, Reinheit und Gehalt nach Einengen im Vakuumzentrifugalkonzentrator photometrisch bestimmt. Zur Konzentrationsberechnung werden die entsprechenden molaren Extinktionskoeffizienten  $\epsilon$  herangezogen. Der molare Extinktionskoeffizient für pC-CoA bei 333 nm beträgt  $21 \times 10^6 \text{ cm}^2/\text{mol}$ , für Caf-CoA ist  $\epsilon$  (346 nm)  $18 \times 10^6 \text{ cm}^2/\text{mol}$  (Stöckigt und Zenk, 1975).

### 3.2.5. Proteinreinigung über His-Tag Affinitätschromatographie

Das verwendete Säulenmaterial zur Aufreinigung ist Ni-NTA His Bind<sup>®</sup> Resin (Novagen). Bei der Aufreinigung von Proteinen mit His-Tag wird die Affinität des His-Tags zu den immobilisierten Nickelionen ausgenutzt (Abb. 3.4).

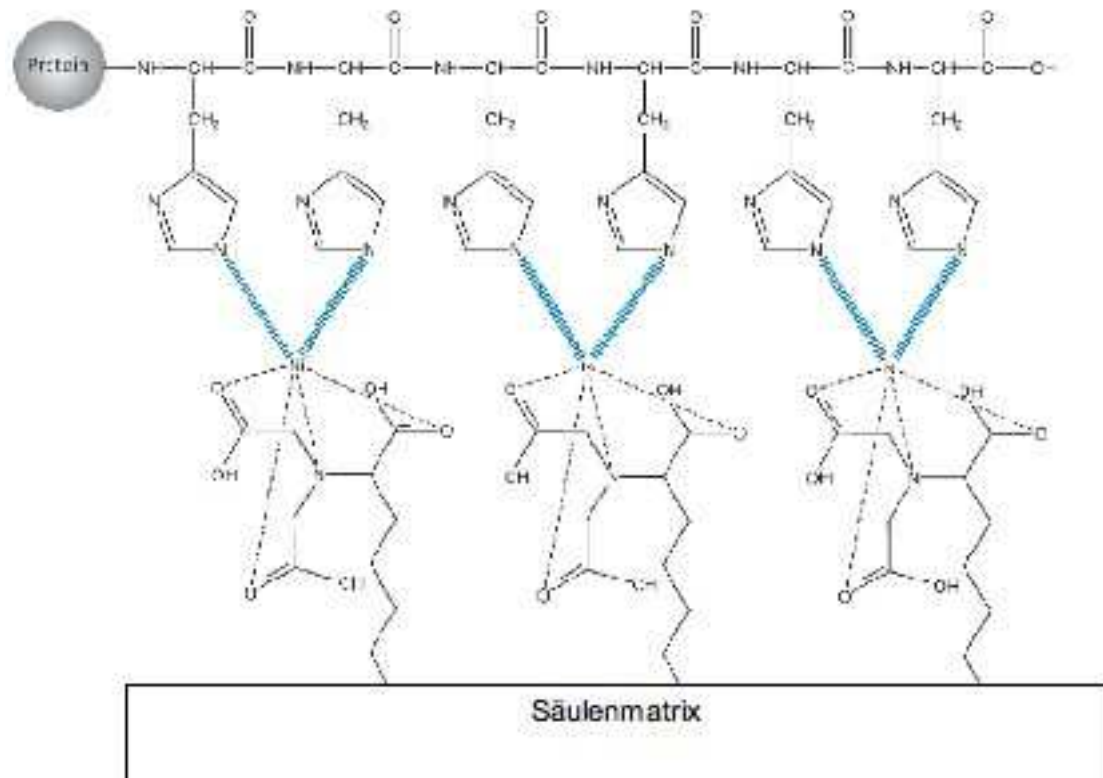


Abb. 3.4: Bindungsmuster eines Hexahistidintags an eine Ni-NTA Matrix (Quelle: [http://www.nature.com/app\\_notes/nmeth/2008/081305/images/Figure1.gif](http://www.nature.com/app_notes/nmeth/2008/081305/images/Figure1.gif). Grafik abgeändert).

Zur Elution wird Imidazol verwendet, welches mit dem His-Tag um die Bindestellen konkurriert. Bei genügend hoher Imidazolkonzentration wird der His-Tag aus der Bindung getrennt und eluiert zusammen mit dem gesuchten Protein von der Säule. Die Bindekapazität des Säulenmaterials wird mit 5-10 mg Protein pro ml Säulenmaterial angegeben.

Zur Durchführung wird 1 ml des Säulenmaterials mit 4 ml Bindepuffer (50 mM Natriumphosphatpuffer pH 8,0, 300 mM NaCl, 10 mM Imidazol) äquilibriert. Danach wird das Säulenmaterial in eine leere gesäuberte PD10-Säule überführt. Der

Äquilibriumspuffer wird entfernt und der Proteinrohextrakt auf die Säule gegeben. Nach Zugabe des Proteinrohextraktes zur Nickelmatrix wird die Säule verschlossen und auf Eis gelagert. Es erfolgt eine Inkubation für 60 min auf einem Schüttler. Das 6xHis-Fusionsprotein bindet während der Inkubation an die Nickelmatrix. Der Proteinrohextrakt wird abgelassen und als Durchlauf gesammelt. Idealerweise ist bei der Kontrolle des Durchlaufs kein heterolog exprimiertes Protein zu detektieren. Zwei Waschschriffe mit jeweils 4 ml Waschpuffer (50 mM Natriumphosphatpuffer pH 8,0, 300 mM NaCl, 20 mM Imidazol) folgen. Der Waschpuffer sollte dabei auch eine Kontaktzeit von wenigen Minuten haben, um Verunreinigungen besser von der Matrix lösen zu können. Das heterolog exprimierte Protein wird mit aufeinanderfolgenden Elutionsschriffen mit je 0,5 ml Elutionspuffer (50 mM Natriumphosphatpuffer pH 8,0, 300 mM NaCl, 250 mM Imidazol) eluiert.

Durchlauf, Waschfraktionen und die Elutionsfraktionen werden getrennt gesammelt. Zur Kontrolle des Reinigungseffekts werden Aktivitätstests, Proteingehaltsbestimmungen und SDS-Gelelektrophorese durchgeführt.

### **3.2.6. Umpuffern der Elutatfraktionen über Größenausschlusschromatographie (SEC)**

Bei der SEC werden kleinere Moleküle stärker retardiert als große. Dadurch ergibt sich ein größenorientiertes Aufreinigungsmuster, das alleine durch die Bewegungseigenschaften der einzelnen Moleküle durch die Säule erzeugt wird. Die Wechselwirkung einzelner Moleküle mit dem Material spielt hier idealisiert keine Rolle. Wichtig ist die Einhaltung von Volumenobergrenzen der verwendeten Probe. Das Volumen der Probe sollte höchstens ein Drittel des Säulenvolumens sein, um eine gute Trennung zu erzielen. Eine optimale Aufreinigung zeigt separierte Peaks, die nicht ineinander laufen.

Zur Durchführung werden PD10-Säulen mit 8,3 ml Sephadex G-25 Matrixgel als stationäre Phase verwendet. Das Ausschlussvolumen des Säulenmaterials liegt bei 5 kDa. Gelöste Substanzen, die diese GröÖe überschreiten, werden zuerst eluiert. Die Elutatfraktionen der His-Tag-Aufreinigung werden wie folgt weiter verarbeitet: Die Säule wird mit 25 ml KPi-Puffer (pH 7,0) äquibriert. In diesen Puffer wird das Protein

umgepuffert. Anschließend wird die Probe auf die Säule pipettiert und einsickern gelassen. Das Probenvolumen sollte 2,5 ml nicht übersteigen. Nun werden die Proteine mit 3,5 ml KPi-Puffer (pH 7,0) eluiert.

### **3.2.7. Umpuffern und Konzentrierung durch Dialyse**

Die Dialyse ist eine schonende Methode, um Proteinlösungen umzupuffern und zu konzentrieren. Die Dialyse basiert auf dem Prinzip des Konzentrationsausgleichs durch osmotisches Potential über eine semipermeable Membran. Hochmolekulare Stoffe werden zurückgehalten und niedermolekulare strömen durch die Membran bis ein Gleichgewicht erreicht ist.

Zur Durchführung wird ein Dialysierschlauch (ZelluTrans Roth®) aus regenerierter Cellulose mit einer Ausschlussgrenze von 5 kDa verwendet. Vor dem Befüllen des Schlauchs mit der Proteinlösung muss die Membran 20 min in destilliertem Wasser konditioniert werden. Danach wird der Schlauch an einer Seite mit einer Klammer verschlossen oder verknotet und die Proteinlösung eingefüllt. Nachdem der Schlauch mit der zweiten Klammer verschlossen ist, wird er in den Dialysepuffer überführt. Salze und Puffersubstanzen mit niedrigem Molekulargewicht wie Phosphatpuffer und Tris/HCl äquilibrieren unter Rühren innerhalb von 3 Stunden (Herstellerangaben). Moleküle mit einem Molekulargewicht von über 5 kDa passieren die Membran nicht. Die in dieser Arbeit untersuchten Moleküle enthalten > 200 Aminosäuren (> 20 kDa) und werden somit zurück gehalten.

Das Volumen des Dialysierpuffers soll mindestens dem 50-fachen Volumen der Probe entsprechen. Es wird mindestens für 3 Stunden im Kühlraum dialysiert. Der Dialysierpuffer wird einmal erneuert und die Probe über Nacht bei 4 °C umgepuffert. Zur Konzentrierung der Proteinlösungen verbleibt die umgepufferte Probe im Dialyseschlauch und wird in SpectraGel™ Absorbent eingebettet. SpectraGel™ Absorbent ist ein wasserunlösliches Polyacrylgranulat, welches hygroskopisch wirkt und der Probe Wasser entzieht. Eine Konzentrierung auf das Zehnfache erfolgt innerhalb von ca. 1 h.

### 3.2.8. Proteingehaltsbestimmung nach Bradford

Die Proteinkonzentration sämtlicher Proteinrohextrakte und gesammelten Fraktionen während der Aufreinigung wird über die Methode nach Bradford (1976) bestimmt. Der Triphenylmethanfarbstoff Coomassie Brilliant Blue G-250 bildet in saurem Milieu unspezifische Komplexe mit kationischen und unpolaren Seitenketten von Proteinen. Dabei verschiebt sich das Absorptionsmaximum von 465 nach 595 nm. Bei dieser Wellenlänge wird die Absorption der Proben gemessen. Durch Vergleich der Absorption der Proben mit derjenigen eines BSA-Standards eingestellt auf 1 mg/ml kann die Proteinkonzentration errechnet werden. Eine Referenz mit Puffer anstelle der Proteinlösung dient als Nullabgleich.

Zur Durchführung wird 2 ml Bradford-Reagenz im Verhältnis von 1:200 gemischt, jeweils mit:

- 10 µl            Puffer als Referenz
- 10 µl            1 mg/ml BSA-Standard
- 10 µl            zu messende Proteinlösung

Diese Ansätze inkubieren 15 Minuten bei Raumtemperatur bevor sie photometrisch gegen die Referenz vermessen werden. Die Zusammensetzung des Bradford-Reagenzes kann Tab. 3.1 entnommen werden.

Tab. 3.1: Zusammensetzung Bradford-Reagenz

|        |                                |
|--------|--------------------------------|
| 100 mg | Coomassie Brilliant Blue G 250 |
| 50 ml  | 96% Ethanol                    |
| 850 ml | destilliertes Wasser           |
| 100 ml | 85% o-Phosphorsäure            |

Zum Entfernen von Schwebeteilchen wird der Ansatz doppelt filtriert.

### 3.2.9. SDS-PAGE

Bei der SDS-PAGE (sodium dodecylsulfate polyacrylamide gel electrophoresis, Natriumlauroylsulfat-Polyacrylamidgelelektrophorese) werden Proteingemische unter denaturierenden Bedingungen entsprechend ihrer Größe elektrophoretisch aufgetrennt. Die Wanderungsgeschwindigkeit von Proteinmolekülen in einem elektrischen Feld ist dabei direkt proportional zu ihrer Gesamtladung und indirekt proportional zu ihrer Masse. Die Nettoladung eines Proteins in wässriger Lösung ist vom pH-Wert der Lösung und dem isoelektrischen Punkt, der sich aus der Aminosäuresequenz und deren  $pK_a$ -Werten ergibt, abhängig. Native Proteine würden deshalb ihrer Nettoladung entsprechend entweder zur Anode oder zur Kathode wandern. Der Einfluss der Nettoladung wird durch SDS unterdrückt, indem es Proteine entfaltet und bindet. Dort erzeugt SDS über seine anionische Kopfgruppe eine hohe negative Gesamtladung im SDS-Protein-Komplex. Die ursprüngliche Nettoladung ist demgegenüber vernachlässigbar. Da die Menge an gebundenem SDS mit der Masse des jeweiligen Proteins korreliert, wird eine Trennung nach Molekulargewicht erreicht.

Ein Marker, der verschiedene Proteine mit bekanntem Molekulargewicht enthält, ermöglicht die Bestimmung der Größe der Proteine. Die Güte der Aufreinigung des Proteins wird ebenso durch dieses Verfahren bestimmt. Nach der His-Tag-Aufreinigung sollte idealerweise nur noch die Proteinbande des überexprimierten Proteins zu finden sein. Nach der Elektrophorese wird das Ergebnis durch Färbung mit Coomassie-Reagenz sichtbar gemacht. Durch die enthaltene Essigsäure im Reagenz werden die Proteine im Gel fixiert. Gleichzeitig lagert sich der Coomassie-Farbstoff an die Proteine an. Anschließend wird das Gel in Entfärbelösung entfärbt. Nur an den Stellen, an denen die Proteine fixiert sind, bleibt die blaue Farbe sichtbar. Das Ergebnis kann fotografisch dokumentiert werden. Die Sensitivität der Coomassie-Färbung liegt bei 50 ng Protein pro Bande. Das Probenvolumen richtet sich nach dem Reinheitsgrad und der Proteinkonzentration.

Die Durchführung wurde von Ulrich Laemmli in abgewandelter Form schon 1970 beschrieben. Dafür werden zunächst die Acrylamidgele gegossen. Die verwendeten Glasplatten müssen gesäubert und die Gelgießapparatur zusammengebaut werden. Der Probenkamm wird kurzzeitig eingesetzt, um die Höhe des Trenngels 1 cm unter

den Taschen zu markieren. Die Zusammensetzung des Trenngels kann Tab. 3.2 entnommen werden.

Tab. 3.2: Zusammensetzung des Trenngels

|         |   |                    |
|---------|---|--------------------|
| 2,5 ml  | 1,5 M Tris/HCl pH 8,0                         |                    |
| 3,35 ml | destilliertes Wasser                          |                    |
| 4,1 ml  | 30% Acrylamid-Stammlösung mit 2% Bisacrylamid |                    |
| 50 µl   | 20% SDS                                       |                    |
| 16 µl   | TEMED   |                    |
| 70 µl   | 10% APS                                       | Zuerst<br>wird das |

Trenngel direkt nach der Zugabe von APS bis zur markierten Höhe zwischen die Glasplatten gegossen. Das Trenngel wird sofort mit 0,5 ml 1-Butanol überschichtet, um Grenzflächeneffekte zwischen Gel und Luft auszuschalten. Nach ca. 20 min ist das Trenngel ausgehärtet. Das Butanol wird mit Filterpapier entfernt. Nun wird das Sammelgel (Tab. 3.3) gemischt und über das Trenngel geschichtet. In das noch flüssige Sammelgel wird der Probenkamm blasenfrei eingesetzt.

Tab. 3.3: Zusammensetzung des Sammelgels

|         |   |
|---------|---|
| 1,25 ml | 0,5 M Tris/HCl pH 6,8                         |
| 2,97 ml | destilliertes Wasser                          |
| 0,75 ml | 30% Acrylamid-Stammlösung mit 2% Bisacrylamid |
| 25 µl   | 20% SDS                                       |
| 10 µl   | TEMED   |
| 35 µl   | 10% APS                                       |

Nach dem Start der Polymerisation durch N,N,N',N'-Tetramethyldiamin (TEMED) und Ammoniumperoxydisulfat (APS) müssen die Gele zügig gegossen werden. Sie können nach dem Aushärten in die Elektrophoreseapparatur eingesetzt werden. Die Elektrophoreseapparatur wird nach Herstellerangaben aufgebaut und mit Elektrophorese-Laufpuffer (Tab. 3.4) bis zur Markierung aufgefüllt.

Tab. 3.4: Zusammensetzung des Laufpuffers

|        |                 |
|--------|-----------------|
| 0,1%   | SDS             |
| 192 mM | Glycin          |
| 25 mM  | Tris/HCl pH 8,3 |

Der Gelkamm wird vorsichtig entfernt, so dass die Taschen intakt bleiben und keine Luftblasen in den Taschen verbleiben.

Die Rohextraktproben werden entweder direkt oder nach Aufkonzentrierung durch TCA-Fällung (siehe 3.2.10 TCA-Fällung) verwendet. Die mit Roti<sup>®</sup>-Load vermischten Proben (3 Teile Probe, 1 Teil Roti<sup>®</sup>-Load) werden zum Denaturieren für 3 bis 5 min in kochendem Wasser erhitzt, kurz zentrifugiert und in die Probenaschen einpipettiert. Eine Probenasche wird mit 10 µl Marker (Roti<sup>®</sup>-Mark STANDARD) befüllt. Nach elektrophoretischer Trennung bei 200 V für 2 bis 3 h wird das Gel vorsichtig aus den Glasplatten entnommen und mit Coomassie-Färbelösung (Tab. 3.5) gefärbt.

Tab. 3.5: Zusammensetzung der Färbelösung

|         |                                |
|---------|--------------------------------|
| 0,625 g | Coomassie Brilliant Blue R 250 |
| 50 ml   | Eisessig                       |
| 225 ml  | destilliertes Wasser           |
| 225 ml  | Methanol                       |

Das Gel wird mindestens 1 h bei Raumtemperatur gefärbt. Eine längere Inkubation ist möglich, da durch das folgende Entfärben des Gels mit Entfärbelösung (Tab. 3.6) unspezifische Bindungen des Coomassie-Farbstoffs an die Gelmatrix ausgewaschen werden.

Tab. 3.6: Zusammensetzung der Entfärbelösung:

|        |                      |
|--------|----------------------|
| 50 ml  | Eisessig             |
| 225 ml | destilliertes Wasser |
| 225 ml | Methanol             |



Das Gel wird bis zum gewünschten Ergebnis in Entfärbelösung geschwenkt. Idealerweise sind nur noch die Proteinbanden klar abgegrenzt voneinander zu erkennen. Die Gelmatrix verliert die Blaufärbung vollständig.

### **3.2.10. TCA-Fällung zur Aufkonzentrierung**

Das Gesamtprotein kann unter denaturierenden Bedingungen mit Trichloressigsäure (TCA) gefällt und somit aufkonzentriert werden.

Zur Durchführung werden zu jeder Probe Trichloressigsäurelösung (40%) zu gleichen Teilen gegeben. Nach Umschwenken erfolgt die Fällung nach ca. 1 h bei 4 °C im Kühlschrank. Das gefällte Gesamtprotein wird durch Zentrifugation bei 4 °C sedimentiert (15000 g, 20 min). Der Überstand wird vorsichtig abgehoben und verworfen. Das Sediment wird zweimal in 500 µl eiskaltem Aceton (80%) gewaschen. Das in der Vakuumzentrifuge getrocknete Sediment wird in 20 µl 0,1 M Natriumcarbonatlösung mit 10% Saccharose gelöst.

### **3.3. *Bestimmung der Enzymaktivität von RAS, HQT und HST***

Die Hydroxycinnamoyl-CoA:Hydroxyphenyllactat Hydroxycinnamoyltransferase (Rosmarinsäuresynthase, RAS), die Hydroxycinnamoyl-CoA:Shikimat Hydroxycinnamoyltransferase (HST) und die Hydroxycinnamoyl-CoA:Chinat Hydroxycinnamoyltransferase (HQT) sind CoA-Ester-abhängige BAHD-Acyltransferasen, die an der Biosynthese von Rosmarinsäure, Chlorogensäure und Monolignolen beteiligt sind (siehe 1. Einleitung). Sie unterscheiden sich hinsichtlich ihrer Substratspezifität. Für die Donatoren besitzen die Enzyme eine moderate und für die Akzeptoren in der Regel eine hohe Substratspezifität. Sie übertragen 4-Cumaroyl-CoA sowie Caffeoyle-CoA und auch andere (Hydroxy-)Cinnamoyl-Reste aus ihren CoA-Thioestern auf ihre Akzeptorsubstrate (Petersen et al. 2009). Die enzymatisch hergestellten Produkte unterscheiden sich durch das Substitutionsmuster an den Aromaten. In dieser Arbeit wird getestet, ob RAS und

HST in weiteren Pflanzen zu finden sind. Die dabei untersuchten enzymatischen Umsetzungen werden in Abb. 3.5 aufgelistet.

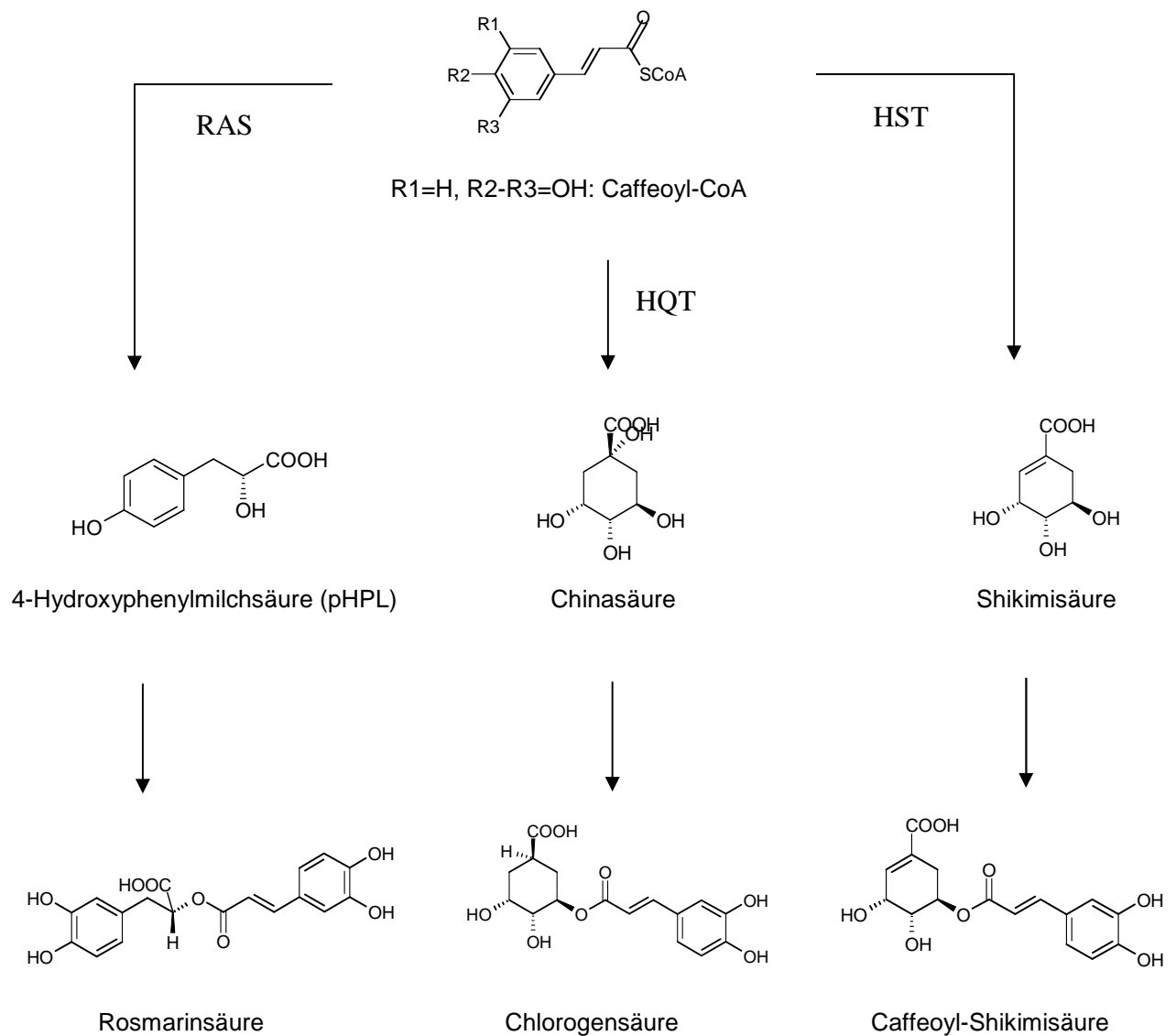


Abb. 3.5: Schema der durchgeführten Enzymtests mit Proteinrohextrakten zum Auffinden der Hydroxycinnamoyltransferasen RAS, HST und HQT. Die Endprodukte werden über HPLC detektiert und zeigen somit indirekt die Umsetzung durch die Enzyme an.

Zu den Pflanzenarten, von denen Gene für Hydroxycinnamoyltransferasen bekannt sind, zählen beispielhaft (in Klammern UniProt Nummer):

|                            |              |
|----------------------------|--------------|
| <i>Coleus blumei</i>       | RAS (A0PDV5) |
| <i>Melissa officinalis</i> | RAS (G12794) |
| <i>Cynara cardunculus</i>  | HQT (A4U7G0) |
| <i>Trifolium pratense</i>  | HCT (B6D7P1) |
| <i>Nicotiana tabacum</i>   | HST (Q8GSM7) |

|                          |              |
|--------------------------|--------------|
| <i>Coffea canephora</i>  | HCT (A4ZKE4) |
| <i>Nicotiana tabacum</i> | HQT (Q70G33) |
| <i>Cynara scolymus</i>   | HQT (C3PTS7) |

Konnten mögliche Sequenzen für RAS, HQT oder HST aus Pflanzen isoliert werden, so wurden diese in spezielle Vektoren kloniert (siehe 3.1. Klonierung von RAS und HST aus Pflanzenmaterial). Diese Vektoren wurden zur Transformation von Bakterien verwendet, mit dem Ziel die Enzyme in Bakteriensuspensionskulturen zu gewinnen. Nach Inkubation der Bakteriensuspensionskulturen wurden diese zu einem Proteinrohextrakt weiter verarbeitet und aufgereinigt (siehe 3.2. Expressionsversuche der gefundenen RAS und HST cDNA).

Der Proteinrohextrakt und sämtliche aufgesammelten Fraktionen aus der Aufreinigung werden auf RAS-, HQT- und HST-Aktivität hin untersucht. Dazu wurden folgende Donor- und Akzeptorsubstratkombinationen getestet:



Der Ansatz wird durch HPLC-Analyse (siehe 3.3.2. HPLC) auf die Bildung der Reaktionsprodukte untersucht. Die spezifischen Aktivitäten können aus den Mengen der Reaktionsprodukte berechnet werden. Um zu gewährleisten, dass die Enzymreaktion im linearen Bereich erfasst wird, werden unverdünnte, 1:10 und 1:100 verdünnte Proteinlösungen in den Enzymtests eingesetzt.

### **3.3.1. Enzymtests mit pHPL, Chinasäure und Shikimisäure**

Für die Versuche zur biochemischen Charakterisierung werden die Rohextrakte wie beschrieben (siehe 3.2.3. Herstellung von Proteinrohextrakt) mit KPi-Puffer hergestellt. Die über His-Tag-Affinitätschromatographie aufgereinigten Eluatfraktionen werden in 20 mM Tris/HCl pH 7,5 mit 0,02% NaN<sub>3</sub> umgepuffert. Die

Kulturansätze werden nicht gemischt verwendet. Zur Kontrolle werden heterolog hergestellte, aufgereinigte Proteine aus anderen Pflanzenarten verwendet. Die CbRAS-Lösung hatte eine Proteinkonzentration von 0,84 mg/ml. Die Proteinkonzentration der CbHST-Lösung betrug 0,34 mg/ml. Die Enzymlösungen werden bei 4 °C gelagert. Von den Akzeptorsubstraten und CoA-aktivierten Säuren (siehe 3.2.4. Synthese der CoA-aktivierten Säure Caf-CoA) werden ebenfalls Stammlösungen angesetzt und bei -20 °C aufbewahrt. Enzyme, Akzeptoren und CoA-aktivierte Säuren werden je nach Versuchsdurchführung in unterschiedlichen Konzentrationen eingesetzt.

Es wurde ein Standardenzymtest zum Überprüfen der Expression der Vollängen-cDNAs etabliert. Alle heterolog exprimierten Enzyme werden auf Enzymaktivität mit den genannten Substraten (pC-CoA, Caf-CoA mit pHPL, Chinat und Shikimat) getestet. Die Durchführung des Standardenzymtests erfolgte nach dem in Tab. 3.7 abgebildeten Pipettierschema.

Tab. 3.7: Pipettierschema des verwendeten Standardenzymtests zur Analyse von Proteinrohextrakten und Eluatfraktionen

|                          |                             |
|--------------------------|-----------------------------|
| 90,5 µl                  | 0,1 M KPi-Puffer            |
| 5 µl                     | 12,5 mM Ascorbat            |
| 12,5 µl                  | 0,01 M DTT                  |
| 5 µl                     | 100 mM Akzeptorsubstrat     |
| 2 µl                     | 1-3 mM CoA-aktivierte Säure |
| 10 µl                    | Proteinrohextrakt           |
|                          | Eluatfraktionen             |
| Gesamtvolumen von 125 µl |                             |

KPi-Puffer:

0,1 M  $\text{KH}_2\text{PO}_4/\text{K}_2\text{HPO}_4$ -Puffer pH 7,0

Die Reaktion wird durch Zugabe des Proteinrohextrakts oder der Eluatfraktion gestartet. Der Ansatz inkubiert für exakt 10 min bei 30 °C. Die Reaktion wird durch Zugabe von 10 µl 6 N HCl abgestoppt, geschüttelt und auf Eis gekühlt.

### 3.3.2. Extraktion der Produkte aus Enzymtests

Zur Extraktion der enzymatisch gebildeten Produkte wird Ethylacetat verwendet.

Zur Durchführung wird der Enzymtest mit 500 µl Ethylacetat kräftig gemischt, 5 Minuten bei 16000 g zentrifugiert und die Ethylacetatphase (oben) in ein neues 1,5 ml Eppendorfgefäß überführt. Dieser Vorgang wird einmal wiederholt. Die Ethylacetatphasen werden vereinigt und eingedampft. Die Rückstände werden in 100 µl Fließmittel für die HPLC rückgelöst, 5 Minuten bei 16000 g zentrifugiert und der HPLC-Analyse unterzogen. Die Reaktionsprodukte aus den RAS-, HST- und HQT-Enzymtests stehen als eingestellte HPLC-Standardlösungen zur Verfügung.

### 3.3.3. Hochleistungsflüssigkeitschromatographie (HPLC)-Analyse

Die Reaktionsprodukte werden durch HPLC-Analyse identifiziert und quantifiziert. Dafür werden die Reaktionsprodukte in den Injektor (20 µl Probenschleife) der HPLC-Anlage injiziert. Das Laufmittel, die mobile Phase, drückt anschließend die Reaktionsprodukte als Teil der mobilen Phase durch die Säule, welche die stationäre Phase enthält. Der Druck, mit dem die mobile Phase durch die Säule gepumpt wird, kann bis zu 200 bar betragen. Dabei werden zwei Arten von Trennprinzipien unterschieden:

- Normal Phase (NP)-HPLC → polare stationäre Phase  
unpolares Lösungsmittel
- Reversed Phase (RP)-HPLC → unpolare stationäre Phase  
polare mobile Phase

In dieser Arbeit wurde ausschließlich mit einem RP-HPLC System gearbeitet.

Zur Durchführung der isokratischen RP-HPLC stehen eine Shandon Hypersil ODS Säule (270x4.0 mm) und eine äquivalente Equisil ODS Säule zur Verfügung. Bei einer Flussrate von 1,0 ml/min wird mit einem Methanol-Wasser-Gemisch, das mit 0,01% Phosphorsäure angesäuert ist, chromatographiert. In Abhängigkeit von der Substratkombination im Enzymtest werden Methanolkonzentrationen zwischen 15 und 60% verwendet. Die Reaktionsprodukte werden spektrophotometrisch bei Wellenlängen zwischen 220 und 343 nm detektiert.

Bei Detektion eines Reaktionsprodukts kann die Berechnung der spezifischen Aktivität des Enzyms erfolgen. Die Enzymaktivität gibt die Umsatzrate eines Enzyms an. Die SI-Einheit der Enzymaktivität ist 1 Katal (kat) und entspricht 1 Mol Substratumsatz pro Sekunde. Eine weitere wichtige Messgröße bei Enzymen ist die spezifische Aktivität. Sie entspricht der Aktivität pro Gesamtproteinmasse.

$$\text{sp. Akt} [\mu_{\text{kat}}/\text{kg}] = \frac{\text{ArP} [\mu\text{l}] \times \text{R} [\mu\text{l}] \times 1000 \times \text{S} [\mu\text{M}]}{\text{ArS} [\mu\text{l}] \times \text{t} [\text{s}] \times \text{E} [\mu\text{l}] \times \text{B} [\text{mg/ml}]}$$

|          |   |
|----------|---|
| sp. Akt. | Spezifische Aktivität   |
| Ar P     | Integrierte Fläche des Produktpeaks bei der HPLC-Analyse      |
| R        | Rücklösungsvolumen der HPLC-Probe                             |
| S        | Konzentration des HPLC-Standards                              |
| Ar S     | Integrierte Fläche des HPLC-Standards                         |
| t        | Zeit  |
| E        | Volumen der Enzymlösung im Enzymtest                          |
| B        | Proteinkonzentration der im Enzymtest verwendeten Enzymlösung |

### **3.4. Statistische und phylogenetische Methoden**

#### **3.4.1. Chi<sup>2</sup>-Test**

Der Chi<sup>2</sup>-Test wird zur Überprüfung von Häufigkeitsverteilungen verwendet (Pearson, 1900). Hierbei werden drei Hypothesentests unterschieden: Verteilungstest, Unabhängigkeitstest und Homogenitätstest. In dieser Arbeit wird der Test auf Homogenität angewendet, um bei unterschiedlichen DNA-Sequenzen die Homogenität der Verteilung der einzelnen Basen zu bewerten.

Ziel ist die Ermittlung des Signifikanzwertes oder auch p-Wert (engl.: p-value). Der p-Wert ist ein Wahrscheinlichkeitswert zwischen 1 und 0. Je größer der p-Wert ist, desto mehr spricht das Ergebnis für die Nullhypothese und somit für die homogene Grundgesamtheit der Daten (Rinne, 2008).

#### **3.4.2. Berechnung der Intronspleißstellen mit NetPlant Gene Server**

Der NetPlantGene Server (<http://www.cbs.dtu.dk/services/NetPGene/>) am Center for Biological Sequence Analysis der Technischen Universität Dänemark ist ein bioinformatisches Werkzeug, um die Übergänge zwischen Exons und Introns in pflanzlichen gDNA-Sequenzen vorherzusagen. Der NetPlantGene Server berechnet mögliche Introns, die an jeder Stelle eines Gens liegen können. Die Position eines Introns vor dem ersten, zweiten oder dritten Nukleotid eines Codons definiert dieses Intron als Intron der Phase 0, 1, oder 2. Dabei kommen Introns vom Typ GT-AG häufig vor. Diese beginnen mit dem Donor Guanin-Thymin und enden mit dem Akzeptor Adenin-Guanin (Hebsgaard et al., 1996).

Introns in HCT-Genen aus *Glechoma hederacea* und *Coleus blumei* werden durch Sequenzvergleich der gDNA von *Arabidopsis thaliana* identifiziert und die Spleißstellen überprüft.



### **3.4.3. BAHD-Acyltransferasen in der pubmed-Datenbank**

Pubmed (<http://www.pubmed.com>) ist eine bioinformatische Datenbank, in der Proteine aller Lebewesen gesammelt werden. In der zentralen Datenbank für Proteinsequenzen kann nach veröffentlichten Proteinsequenzen gesucht werden, die anhand ihrer Akzessionsnummer eindeutig identifizierbar sind.

In der Datenbank sind eindeutig charakterisierte Proteine, aber auch Sequenzen ohne biochemische Charakterisierung oder Sequenzabschnitte hinterlegt (meist gekennzeichnet durch den Zusatz ‚putatives Protein‘). Zu jedem eindeutig bestimmten Protein sind Informationen über die biologische Funktion sowie die katalytische Aktivität zu finden. Weiterhin sind Links zu anderen Datenbanken und relevanten Literaturquellen angegeben.

In dieser Arbeit werden zur phylogenetischen Analyse nur eindeutig charakterisierte BAHD-Acyltransferasen aus Samenpflanzen verwendet. Putative Sequenzen wurden ausgeschlossen.

### **3.4.4. Phylogenetische Analyse von BAHD-Acyltransferasen mit MEGA 5.0**

MEGA 5.0 (Molecular Evolutionary Genetics Analysis) ist eine Software zur phylogenetischen Analyse von DNA- und Proteinsequenzen. Durch statistische Berechnungen, deren Grundlage die Ähnlichkeit der eingesetzten Sequenzen ist, werden phylogenetische Zusammenhänge rekonstruiert. Diese Zusammenhänge erlauben eine Aussage über den Verwandtschaftsgrad der Sequenzen und können als phylogenetischer Stammbaum dargestellt werden (Tamura et al., 2007).

MEGA 5.0 kann verschiedene phylogenetische Analysen und Darstellungen durchführen. Im Folgenden sind drei momentan wichtige Methoden erläutert (Doyle und Gaut, 2000):

- Distanzmethoden

Bei der Distanzanalyse (z.B. Neighbour-Joining-Verfahren (Saitou und Nei, 1987)) wird eine Matrix erstellt, aus der sich Stammbäume erzeugen lassen, die

die evolutionären Änderungen auf Grundlage der Distanzen zwischen den Sequenzen abbilden. Die Idee der Distanzmethoden beruht darauf, die Anzahl gemeinsamer bzw. abweichender Merkmale zu bestimmen, daraus ein Ähnlichkeitsmaß zu berechnen und die Distanzwerte zu einem Baum zu verrechnen (Fitch und Margoliash, 1967). Die Qualität der Merkmale bleibt dabei unberücksichtigt. Es spielt also keine Rolle, ob gemeinsame analoge, symplesiomorphe (keine engere Verwandtschaft) oder synapomorphe (engere Verwandtschaft) Merkmale gezählt werden. Der Nachteil der Methode ist, dass die aus Basensubstitutionen resultierenden Merkmalsmutationen nicht differenziert berechnet werden. Sie werden zu einem einzigen Distanzwert reduziert und es entstehen gleiche Distanzwerte, obwohl die zugrunde liegenden Sequenzen verschieden sind (Saitou und Nei, 1987).

- Maximale Parsimonie

Die kladistische Methode der maximalen Parsimonie (MP) beruht auf dem theoretischen Parsimonieprinzip (Farris, 1970). Das Parsimonieprinzip besagt, dass die einfachste Lösung eines Problems als die Beste anzunehmen ist. Der ausgewählte Sequenzdatensatz wird dazu in homologe, Parsimonie informativ, und in Parsimonie nicht informative Daten unterteilt. Parsimonie informativ sind Daten, die gleiche Nukleotide oder Aminosäuren an einem Locus in einigen aber nicht in allen Sequenzen besitzen. Danach konstruiert das Verfahren eine Vielzahl möglicher Baumtopologien. Von diesen werden die kürzesten Stammbäume ausgewählt und zu einem Konsensusstammbaum verrechnet. Bei hochvariablen Sequenzen kann die Methode nicht mehr angewendet werden, weil sehr lange Astabschnitte im Stammbaum gebildet werden (Jin und Nei, 1990). Die Berechnung dieser Regionen ergibt nicht immer die korrekte evolutionäre Stammbaumtopologie (Felsenstein, 1978).

- Maximale Wahrscheinlichkeit

Die Maximum-Likelihood-Methode (maximale Wahrscheinlichkeit) bezeichnet ein stochastisches Schätzverfahren. Der Unterschied zur maximalen Parsimonie besteht darin, dass die Stammbaumtopologie ausgewählt wird, die die Sequenzvariabilität eines Datensatzes mit der höchsten Wahrscheinlichkeit unterstützt (Felsenstein, 1981). Es wird das Evolutionsmodell gewählt, welches bei einer möglichst geringen Anzahl variabler Parameter einen möglichst hohen log-likelihood-Wert erzielt. Mit dem log-likelihood-Wert wird die Wahrscheinlichkeit der berechneten Stammbaumtopologie angegeben. Um die Berechnung des wahrscheinlichsten Baumes zu erleichtern, werden meist Daten einer vorherigen MP-Analyse als Startbaum verwendet.

## 4. Ergebnisse

### 4.1. *Isolierung von RAS-ähnlichen Teilsequenzen aus verschiedenen Pflanzenarten*

Um Sequenzen mit Homologie zu RAS aus Pflanzenarten der Lamiaceae und Boraginaceae zu amplifizieren, wurden Primer entwickelt, die an die den stark konservierten Motiven DFGWG und HxxxDG entsprechenden Nukleotidsequenzen binden.

#### RAS-Fn7 forward:

5´-GCA AAC GAG CAC CAC TTA TCC GAC GGC GTC GCC GCC-3´  
Translation: ANEHHLSDGVA

#### RAS-Rn6 reverse:

5´-CAC CGC CTT CCC CCA CCC (AG)AA (AG)TC CGG CTC (AG)TA-3´  
Translation: YEPDFGWGKAV

Das amplifizierte Teilstück liegt außerhalb der konservierten Stelle, in der bei vielen Hydroxycinnamoyltransferasen ein Intron („Q-Intron“; St. Pierre und De Luca, 2000) eingefügt ist, und sollte dementsprechend eine Länge von etwa 700 Basenpaaren ergeben. Die zugrunde gelegte RAS-Sequenz aus *Coleus blumei* wurde zuvor in unserem Labor ermittelt (Berger et al., 2006). Die Teilsequenzen wurden durch PCR amplifiziert. Sie zeigen große Ähnlichkeiten zu der Referenzsequenz der RAS aus *Coleus blumei* (siehe Tabelle 4.1). Als Template wurde aus Pflanzengewebe (meist Blattmaterial) präparierte cDNA, RACE-cDNA oder gDNA verwendet.

Die durch PCR amplifizierten Teilstücke wurden in einen pGEM<sup>®</sup>-T Vektor ligiert, der in *E. coli* (JM109 oder DH5α) eingebracht wurde. Die inserierte Nukleotidsequenz wurde sequenziert. Die Bakterien wurden als Glycerol-Dauerkulturen bei -80°C eingefroren.

Im Folgenden wird beispielhaft die Gewinnung der Teilsequenz aus *Lavandula angustifolia* beschrieben:

Die genomische Sequenz der RAS-ähnlichen Teilsequenz aus *Lavandula angustifolia* konnte durch PCR mit gDNA-Template isoliert werden. Die gDNA wurde aus Pflanzengewebe von *Lavandula angustifolia* gewonnen (siehe 3.1.1. Isolierung von genomischer DNA aus Pflanzenmaterial) und im Agarosegel nachgewiesen (Abbildung 4.1).

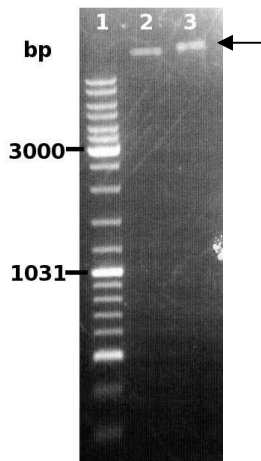


Abb. 4.1: Nachweis isolierter genomischer DNA aus *Lavandula angustifolia* (Bahnen 2 + 3) nach Agarose-Gelelektrophorese, Gel gefärbt mit Ethidiumbromid. Bahn 1: DNA-Größenstandard.

Anschließend wurde die gDNA als Template in einer anschließenden PCR-Reaktion verwendet. Die PCR mit den degenerierten Primern ergab eine Bande mit der Größe von ca. 850 Basenpaaren (Abbildung 4.2).

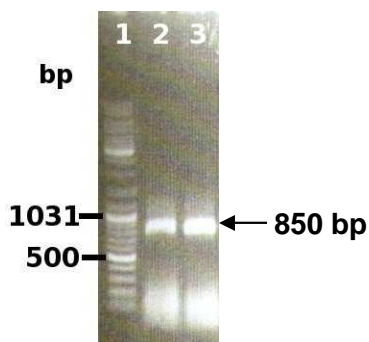


Abb. 4.2: Nachweis eines ca. 850 bp langen PCR-Amplifikats (Bahnen 2 und 3) aus gDNA von *Lavandula angustifolia* mit den Primern Fn7 und Rn6 nach Agarose-Gelelektrophorese, Gel gefärbt mit Ethidiumbromid. Bahn 1: DNA-Größenstandard.

Die ca. 850 Basenpaare lange Teilsequenz wurde aus dem Agarosegel extrahiert, aufgereinigt und in den pGEM<sup>®</sup>-T Vektor ligiert (Abbildung 4.3.).

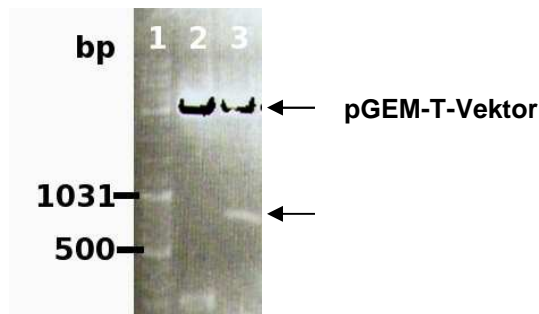


Abb. 4.3: Nachweis der ca. 850 bp großen Sequenz nach Ligation in pGEM<sup>®</sup>-T Vektor (Bahn 2 + 3), Transformation in *E. coli* der Linie DH5 $\alpha$  und Verdau der Plasmidpräparation mit *Nco*I und *Pst*I. Bahn 1: DNA-Größenstandard. Vektor von Bahn 3 enthält das gesuchte Insert und wurde sequenziert.

Nach Transformation von *E. coli* der Linie DH5 $\alpha$  und Plasmidpräparation aus der Übernachtskultur konnte das Plasmid zum Sequenzieren geschickt werden. Die Sequenzanalyse der isolierten Teilsequenz ergab eine Länge von 873 Basenpaaren (siehe 4.2.7. RAS-ähnliche Teilsequenz aus *Lavandula angustifolia*). In diesem Beispiel zeigt sich eine Abweichung von den aus anderen Pflanzenarten amplifizierten Sequenzen, welche eine Länge von ca. 700 bp aufweisen. Die Sequenz von *Lavandula angustifolia* ist mit 873 bp länger als alle anderen amplifizierten putativen RAS-Sequenzen.

Die hohe Ähnlichkeit (siehe Tabelle 4.1) aller gefundenen Teilsequenzen zur RAS aus *Coleus blumei* lässt vermuten, dass die Teilsequenzen tatsächlich RAS-Sequenzen sind, die je nach Pflanzenart etwas variieren können. Es wurde eingehend versucht, das vollständige offene Leseraster der gefundenen Teilsequenzen verschiedener Pflanzenarten mittels RACE-PCR zu finden. Leider konnte keine RAS-Teilsequenz zu einer Vollängensequenz ergänzt werden. Alle hier amplifizierten Nukleotidsequenzen sind im Anhang dieser Arbeit einzeln aufgelistet.

Die gewonnenen Teilsequenzen entsprechen auf Proteinebene zu >74 % der Referenzsequenz einer RAS aus *Coleus blumei*. Die gezeigte Sequenz liegt in dem

Zwischenbereich der stark konservierten Motive HxxxDG und DFGWG (Abbildung 4.4). Die Analyse der erhaltenen DNA-Sequenzen mittels Alignment zeigt deutlich stark konservierte Bereiche (Abbildung 4.5 und 4.6). Die Alignments wurden mit dem CLC Sequence Viewer 6.6.1 berechnet.

Tabelle 4.1: Auflistung der Pflanzen, aus denen vermeintliche RAS-Teilstücke gewonnen werden konnten. In der letzten Spalte ist die Übereinstimmung der gefundenen Sequenzen zur RAS-Referenzsequenz aus *Coleus blumei* (Berger et al., 2006) in Prozent angegeben. Die hier erwähnte Teilsequenz aus *Coleus blumei* ist eventuell eine zweite, bisher unbekannte RAS. Für *Ajuga reptans* wurden zwei Amplifikate isoliert, die mit der Kennung *Ajuga reptans* (1) und (2) versehen sind.

| Familie   | Unterfamilie | Art   | Übereinstimmung<br>DNA-Ebene | Übereinstimmung<br>Proteinebene |
|-----------|--------------|---|------------------------------|---------------------------------|
| Lamiaceae | Ajugoideae   | <i>Ajuga reptans</i>  | (1) 65 % / (2) 81 %          | (1) 50 % / (2) 93 %             |
|           |              | <i>Teucrium chamaedrys</i>  | 96 %                         | 99 %                            |
|           | Lamioideae   | <i>Lamium album</i>   | 99 %                         | 100 %                           |
|           |              | <i>Marrubium vulgare</i>  | 79 %                         | 92 %                            |
|           | Nepetoideae  | <i>Coleus blumei</i> (syn.<br><i>Solenostemon scutellarioides</i> ) | 82 %                         | 93 %                            |
|           |              | <i>Lavandula angustifolia</i>                                       | 81 %                         | 95 %                            |
|           |              | <i>Mentha spicata</i>   | 84 %                         | 93 %                            |
|           |              | <i>Mentha X piperita</i>  | 99 %                         | 97 %                            |
|           |              | <i>Nepeta granadensis</i>   | 82 %                         | 93 %                            |
|           |              | <i>Plectranthus saccatus</i>  | 93 %                         | 97 %                            |
|           |              | <i>Rosmarinus officinalis</i>                                       | 84 %                         | 94 %                            |
|           |              | <i>Salvia farinacea</i>   | 81 %                         | 93 %                            |

#### 4.1.1. Alignment der Aminosäuresequenzen der amplifizierten RAS-Teilstücke

Die Analyse der erhaltenen Aminosäuresequenzen zeigt deutlich stark konservierte Bereiche (Abb. 4.5).

Die gefundenen Teilsequenzen (ca. 210 AS) liegen im Bereich zwischen den stark konservierten Motiven HxxxDG und DFGWG. Der verwendete 5'-Primer enthält die konservierte Sequenz HxxxDG und der 3'-Primer die konservierte DFGWG Sequenz. Die durch die Primer bestimmten Sequenzmotive sind im Vergleich der Aminosäuresequenzen nicht mit einbezogen. In Abb. 4.4 ist der gewählte Bereich innerhalb der RAS-Sequenz von *Coleus blumei* am Beispiel *Nepeta granadensis* dargestellt.

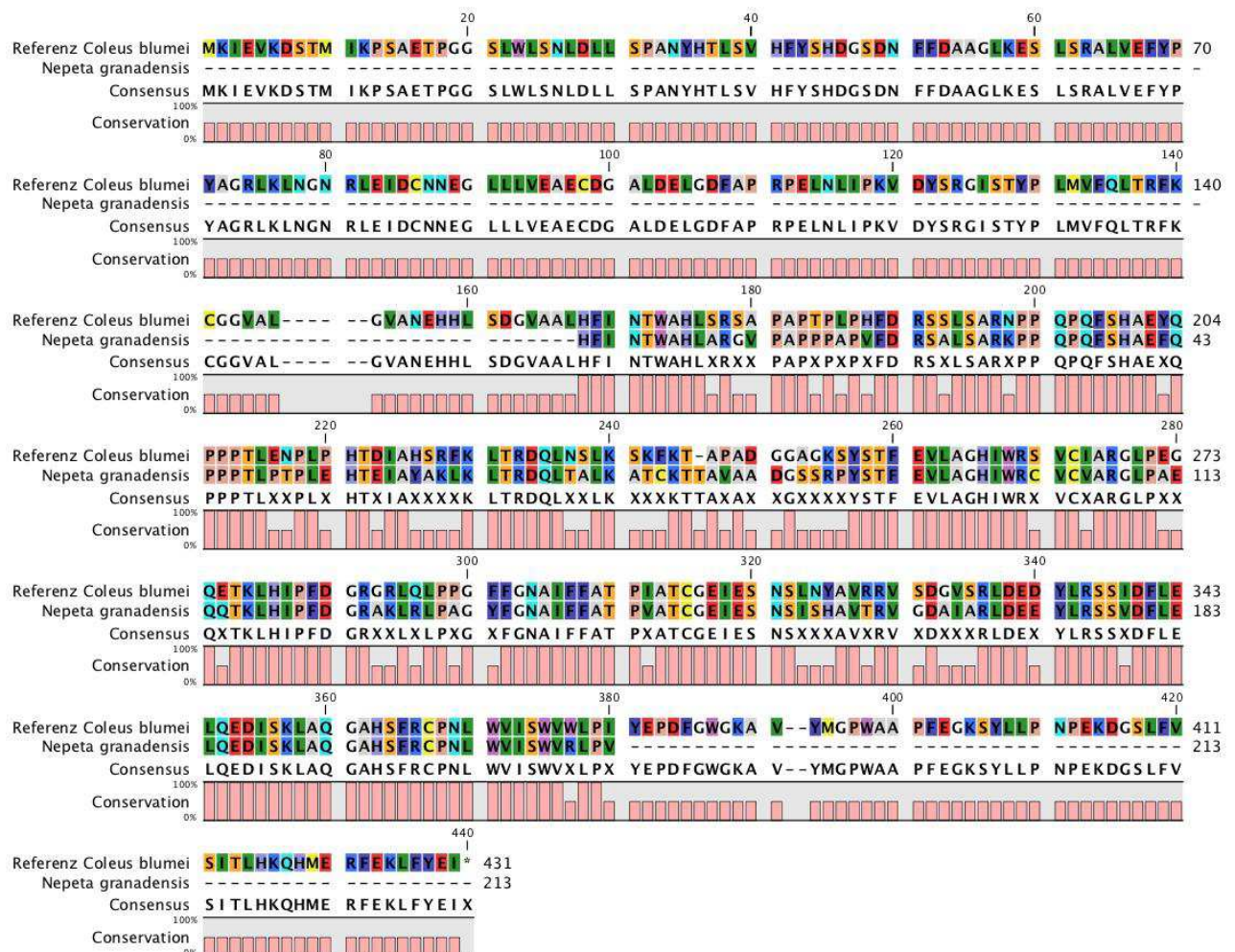


Abb. 4.4: Vergleich der Gesamtsequenz einer RAS aus *Coleus blumei* (Berger et al., 2006) mit dem gefundenen Teilstück zwischen den konservierten Sequenzen HxxxDG und



DFGWG. Beispielhaft wird hier eine putative RAS-Teilsequenz von *Nepeta granadensis* gezeigt.

Die Sequenzlänge zwischen den konservierten Sequenzen HxxxDG und DFGWG bei der RAS aus *Coleus blumei* als Referenz beträgt ungefähr 200 AS. Die gefundenen Teilstücke aus verschiedenen Pflanzenarten sind in Abbildung 4.5 gezeigt. Die Teilstücke besitzen alle die ungefähr erwartete Länge von 200 AS. Die Sequenz von *Lavandula angustifolia* ist mit 241 bp um ca. 20 % länger als die übrigen Sequenzen. Die Länge dieses Zwischenstücks stimmt mit Ergebnissen von Landmann et al. (2011) überein, die auch die RAS und deren Aktivität in *Lavandula angustifolia* nachweisen konnten.

Stark konservierte Bereiche werden im Alignment nochmals hervorgehoben (Abbildung 4.5). Besonders interessant ist die Sequenz HIPFDG (Position AS 152 – 157), welche in allen Sequenzen (außer *Ajuga reptans* 2) zu finden ist. Diese Sequenz enthält dasselbe konservierte Motiv wie die bekannte stark konservierte Sequenz HxxxDG im Bereich des 5'-Primers.

Zusätzlich sind in fast jeder Sequenz (außer bei *Ajuga reptans* 1) folgende Aminosäureabfolgen zu finden:

Position AS 123:   STFVLAGHIWR

Position AS 170:   FGNAIFFATP

Position AS 228:   AQGAHSFRCPNLWVISWV

Die größten Unterschiede in den gefundenen Sequenzen zeigen *Ajuga reptans* 1 und *Mentha X piperita*. Die übrigen Sequenzen sind relativ identisch, besonders wenn Aminosäuren mit gleichen chemischen Eigenschaften (sauer, basisch, aromatisch etc.) als funktional gleichwertiger Austausch betrachtet werden. Aminosäuren mit gleichen chemischen Eigenschaften werden im Alignment mit der gleichen Farbe hinterlegt (Beispiel: basische Aminosäuren sind blau hinterlegt).

Die Primer werden im Alignment nicht mit angegeben. Das Alignment startet direkt nach dem 5'-Primer und endet vor dem 3'-Primer.





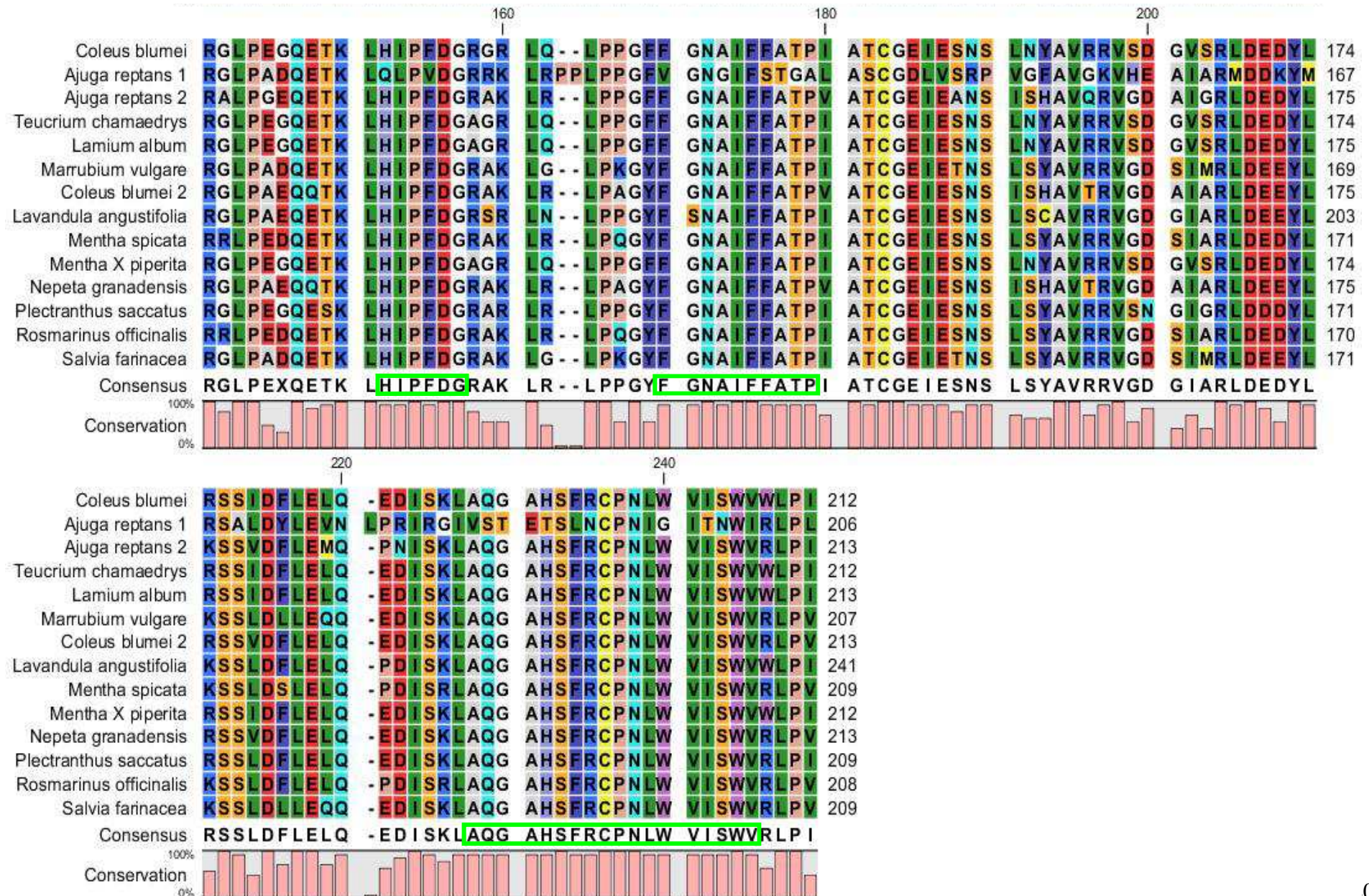


Abb. 4.5: Aminosäuresequenzalignment der amplifizierten Teilstücke einer putativen RAS aus verschiedenen Arten der Lamiaceae. Als Referenz wird in der ersten Zeile die RAS-Sequenz aus *Coleus blumei* gezeigt (Berger et al., 2006). Im Weiteren sind die Pflanzen nach Verwandtschaftsverhältnissen geordnet (siehe Tabelle 4.1). Die am stärksten konservierten Abschnitte sind mit einem grünen Rahmen in der untersten Zeile „Consensus“ markiert.

#### 4.1.2. Alignment der DNA-Sequenzen der amplifizierten RAS-Teilstücke

Mit den degenerierten Primern RAS-Fn7 (forward) und RAS-Rn6 (reverse) konnten aus zuvor hergestellten cDNA und gDNA-Präparationen verschiedener Pflanzenarten (Tabelle 4.1) ca. 700 bp große PCR-Amplifikate isoliert werden (Abbildung 4.6).

Die zuvor gezeigten korrespondierenden Aminosäuresequenzen sind übersichtlicher und zeigen meist größere Gemeinsamkeiten (siehe Tabelle 4.1). Diese größere Übereinstimmung resultiert aus dem Basentriplett: Sind einzelne Nukleotide ausgetauscht, hat dies nicht zwingend einen Austausch der Aminosäure zur Folge.

Die Längenvariabilität der Teilstücke beträgt 13,1%. Dabei ist zu beachten, dass die Sequenz von *Lavandula angustifolia* mit 729 bp wesentlich größer ist als die restlichen Sequenzen. Wird *Lavandula angustifolia* nicht mit betrachtet, reduziert sich die Längenvariabilität auf 3,3%.

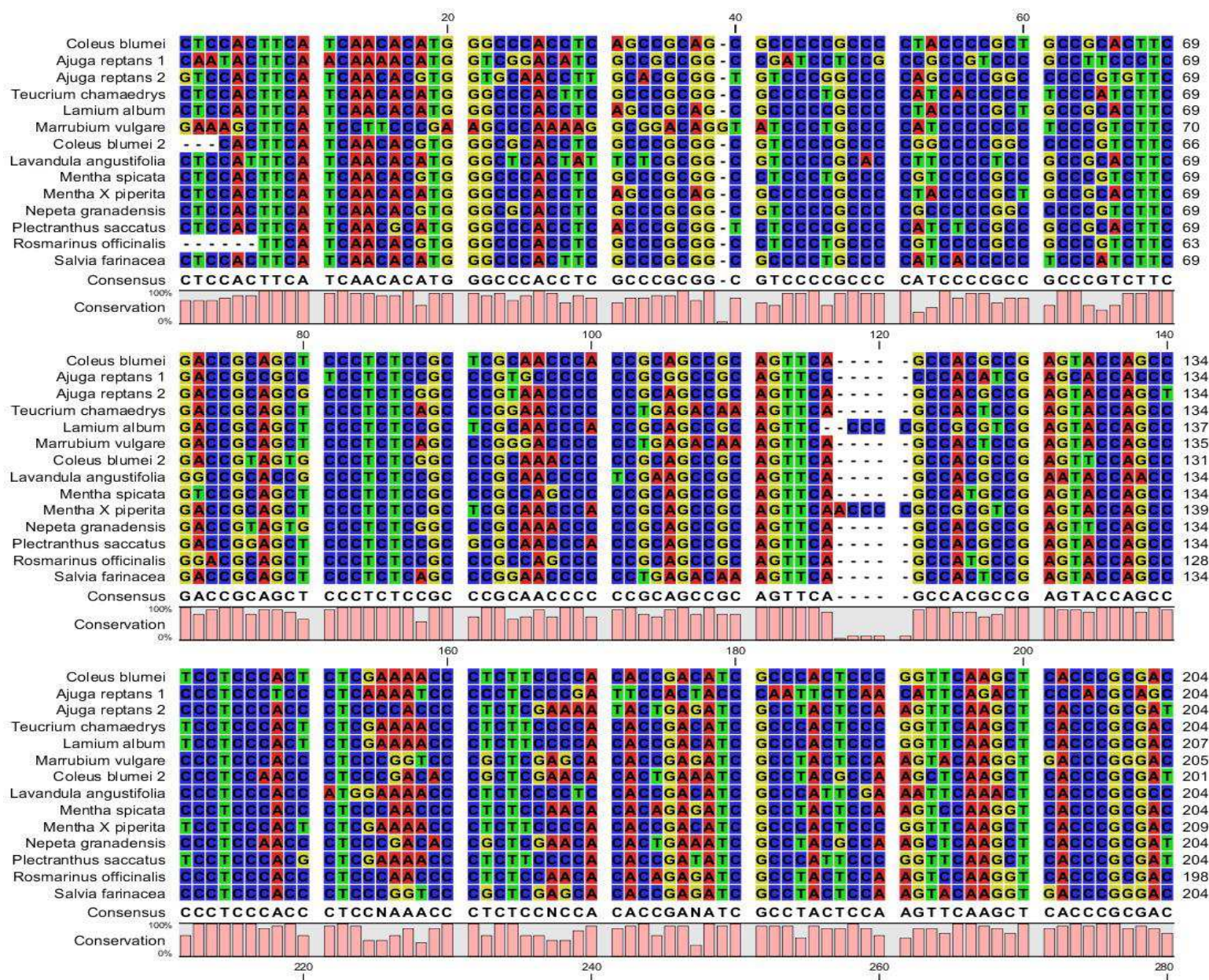
Der  $\chi^2$ -Test (Tabelle 4.2) zeigt die Homogenität in der Basenzusammensetzung der amplifizierten RAS-Teilstücke (p-value = 1,00).

Tabelle 4.2: Basenzusammensetzung der der DNA-Sequenzen der amplifizierten RAS-Teilstücke aus verschiedenen Pflanzen, nach Unterfamilienzugehörigkeit geordnet (siehe Tabelle 4.1). In der Spalte Länge sind die einzelnen Sequenzlängen des kompletten Alignments (ohne die eingesetzten Primer) angegeben. Insgesamt weist der Datensatz eine deutlich homogene Basenzusammensetzung auf (p-value = 1,00). Die statistische Auswertung des Datensatzes erfolgte mit Microsoft Excel.

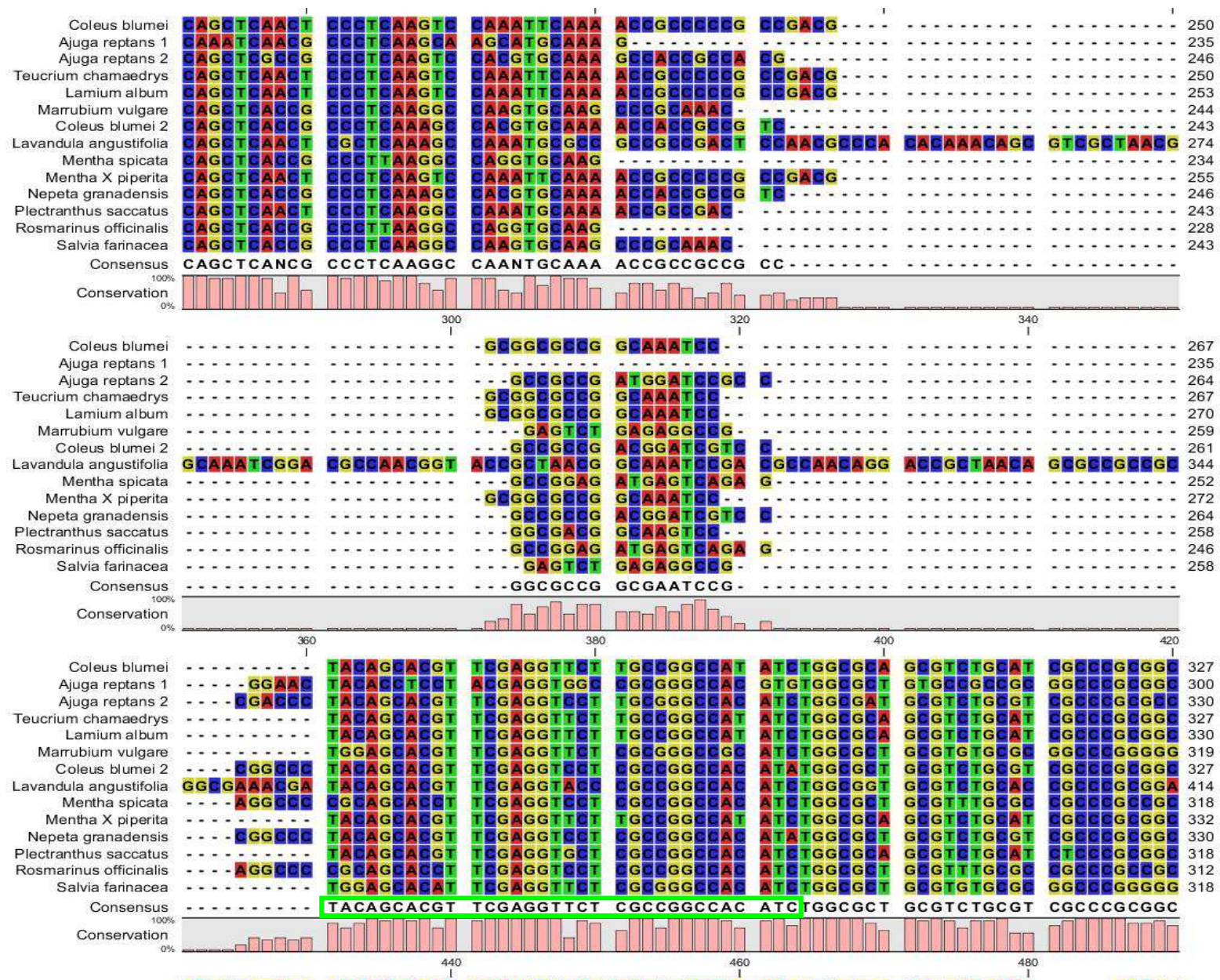
| Pflanze                       | Adenin (A) | Cytosin (C) | Guanin (G) | Thymin (T) | Länge |
|-------------------------------|------------|-------------|------------|------------|-------|
| <i>Ajuga reptans</i> 1        | 0,205      | 0,326       | 0,278      | 0,191      | 621   |
| <i>Ajuga reptans</i> 2        | 0,199      | 0,316       | 0,277      | 0,208      | 642   |
| <i>Teucrium chamaedrys</i>    | 0,191      | 0,377       | 0,249      | 0,183      | 639   |
| <i>Lamium album</i>           | 0,200      | 0,331       | 0,269      | 0,200      | 642   |
| <i>Marrubium vulgare</i>      | 0,217      | 0,287       | 0,305      | 0,191      | 631   |
| <i>Coleus blumei</i> 2        | 0,178      | 0,393       | 0,282      | 0,148      | 639   |
| <i>Lavandula angustifolia</i> | 0,220      | 0,346       | 0,255      | 0,180      | 726   |
| <i>Mentha spicata</i>         | 0,191      | 0,328       | 0,283      | 0,198      | 630   |
| <i>Mentha X piperita</i>      | 0,202      | 0,330       | 0,268      | 0,200      | 644   |
| <i>Nepeta granadensis</i>     | 0,202      | 0,328       | 0,277      | 0,193      | 642   |
| <i>Plectranthus saccatus</i>  | 0,207      | 0,326       | 0,269      | 0,198      | 630   |
| <i>Rosmarinus officinalis</i> | 0,193      | 0,323       | 0,282      | 0,202      | 624   |
| <i>Salvia farinacea</i>       | 0,213      | 0,310       | 0,291      | 0,185      | 630   |
| Mittelwert                    | 0,191      | 0,326       | 0,277      | 0,191      | 642   |

Zur besseren Übersichtlichkeit sind in dem folgenden DNA-Alignment die gleichen Teilbereiche markiert wie in dem Proteinalignment aus Abbildung 4.5.

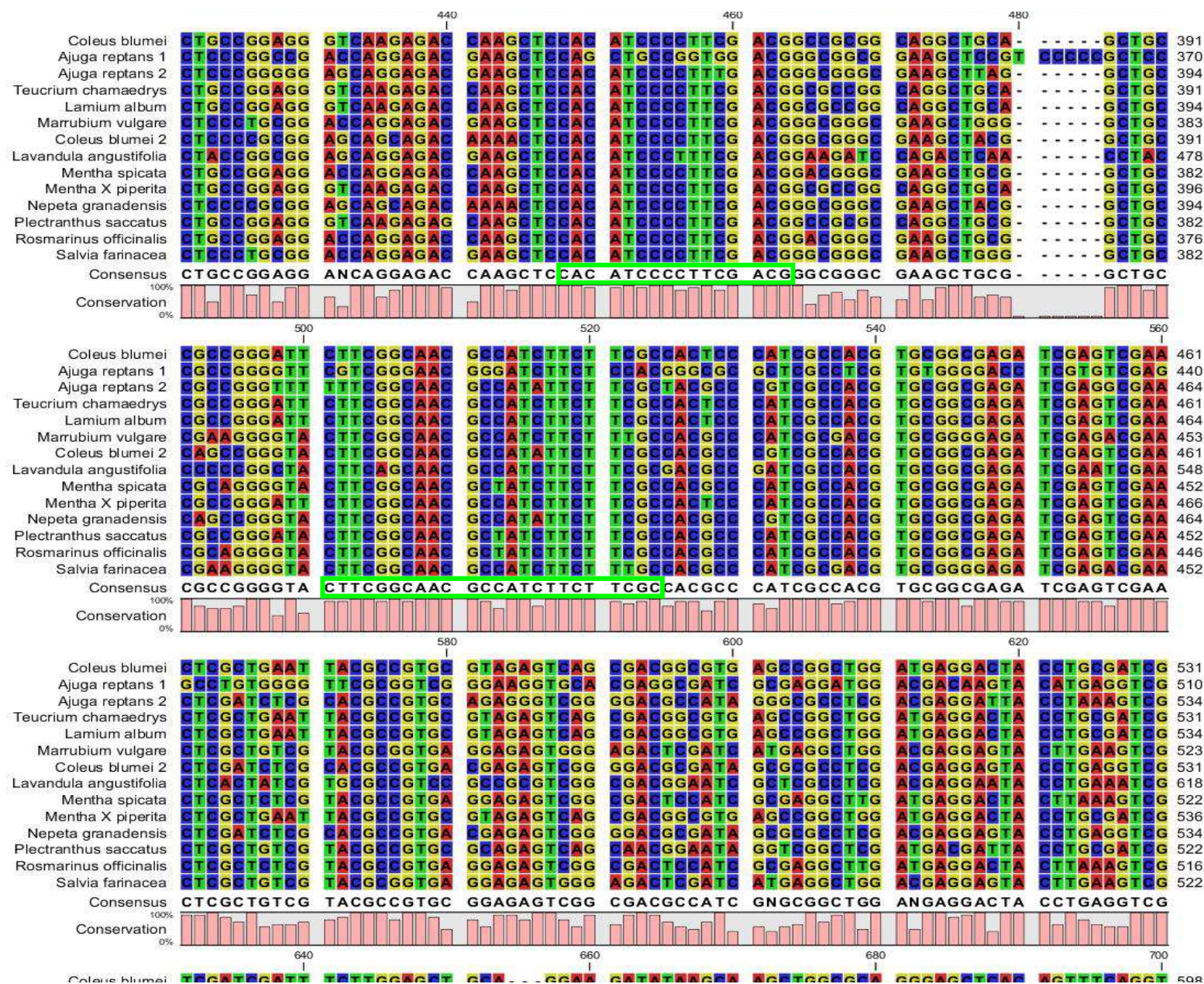














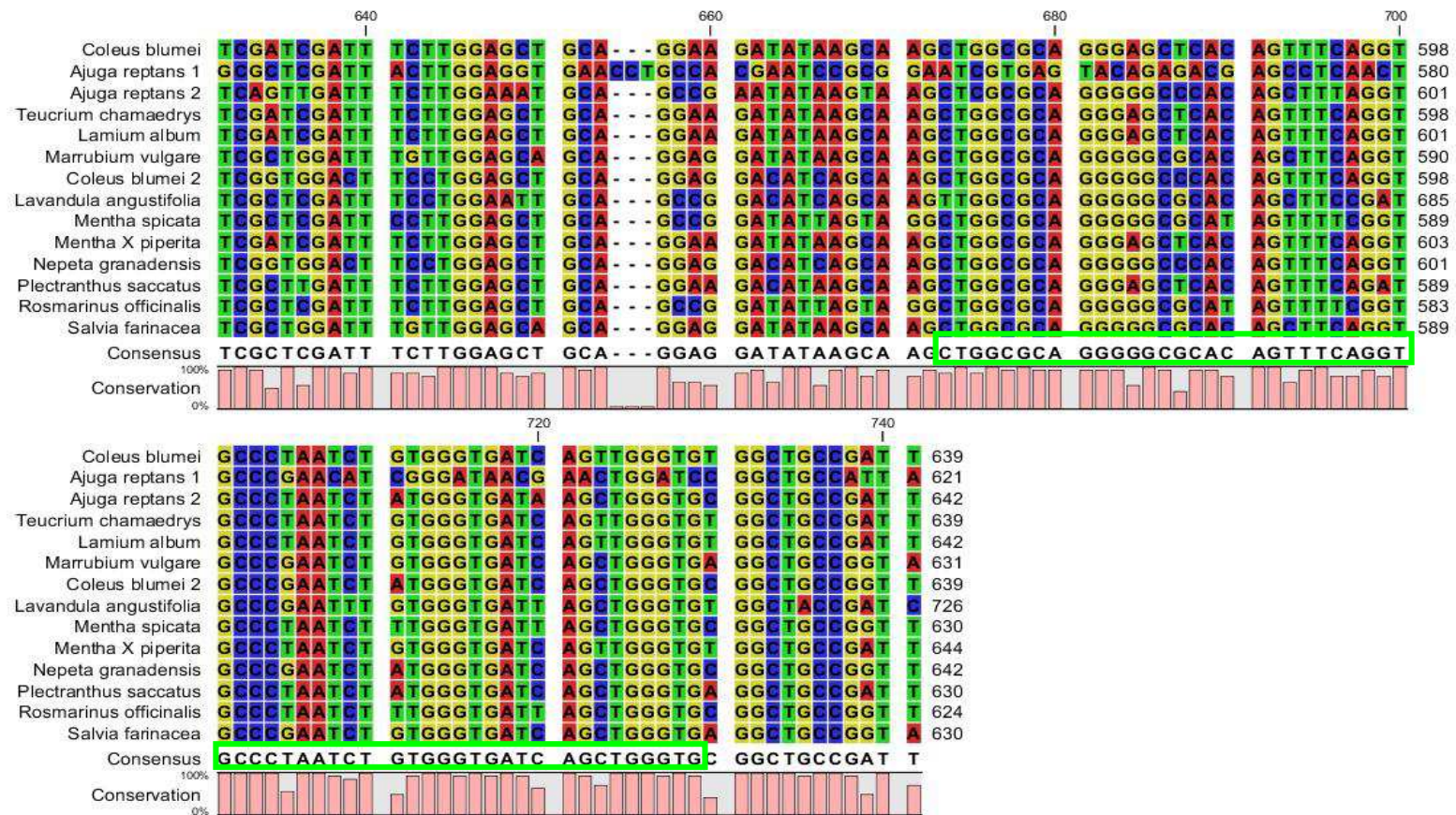


Abb. 4.6: DNA-Sequenzalignment der amplifizierten Teilstücke einer putativen RAS aus verschiedenen Arten der Lamiaceae (Pflanzen nach Verwandtschaftsverhältnissen geordnet). Als Referenz wird in der ersten Zeile die RAS aus *Coleus blumei* gezeigt (Berger et al., 2006). Die am stärksten konservierten Abschnitte sind mit einem grünen Rahmen in der untersten Zeile „Consensus“ markiert und entsprechen den markierten Bereichen des Proteinalignments aus Abbildung 4.5.

## **4.2. Isolierung von HST-ähnlichen Teilsequenzen aus verschiedenen Pflanzenarten**

Bei der Auswertung der erhaltenen putativen RAS-Sequenzen wurde deutlich, dass einige Sequenzen eher der Sequenz von HST-Enzymen entsprachen. Da sich RAS und HST-Enzyme vom Aufbau sehr ähnlich sind (siehe Kapitel 1. Einleitung), war diese Entwicklung plausibel, und eine weitere Strategie zur Gewinnung von potentiellen HST-Sequenzen wurde entwickelt.

Für die Suche nach HST-Genen in verschiedenen Pflanzen wurden spezielle Primer entwickelt. Aus der Erfahrung bei der Suche nach RAS-Genen wurden die zuvor verwendeten RAS-Primer als Grundlage verwendet und zum Auffinden von HST-Genen optimiert. Mit folgendem Primerpaar konnten die besten Ergebnisse erzielt und insgesamt 21 HST-ähnliche Sequenzen amplifiziert werden (siehe Tabelle 4.2):

### HST700f forward:

5'-GGV ATG CAR CAY CAY GYD GCM GAY GGA TTT TCY GGB CTT-3'

Translation: GMQHHXADGFSGL

### HST700r reverse:

5'-AAK ATM GGM CKM CCC CVC CAA ART CWG CAT CAT-3'

Translation: DADFGGXXXX

Als Template wurde aus Pflanzengewebe präparierte cDNA, RACE-cDNA oder gDNA verwendet.

Die Vorgehensweise zur Gewinnung der Sequenzen ist identisch mit der bereits beschriebenen Methode zur Gewinnung der RAS-Sequenzen (siehe Kapitel 4.1 Entdeckung von RAS-ähnlichen Teilsequenzen in verschiedenen Pflanzenarten).

Beispielhaft wird das Agarosegel einer erfolgreichen Amplifizierung von putativen HST-Teilstücken von verschiedenen Pflanzenarten gezeigt (Abbildung 4.7). Die Teilstücke wurden jeweils aus zuvor frisch hergestellter gDNA der jeweiligen Pflanze gewonnen.

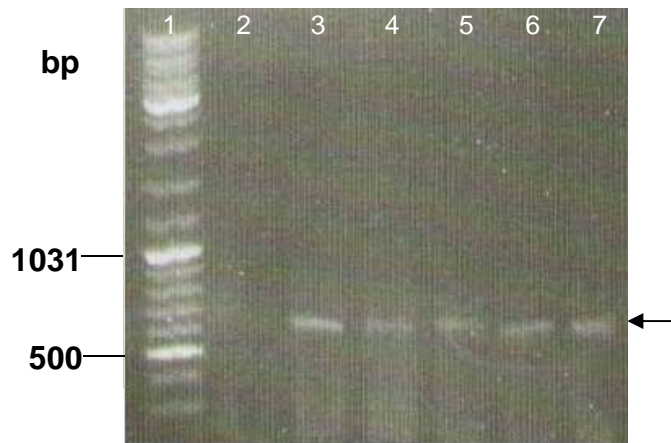


Abb. 4.7: Nachweis eines ca. 650 bp langen PCR-Amplifikats (mit gDNA von Bahn 3: *Coleus blumei*, Bahn 4: *Hyssopus officinalis*, Bahn 5: *Moltkia petraea*, Bahn 6: *Alkanna tuberculata*, Bahn 7: *Onosma angustissima*). Zur Amplifikation wurden die Primer HST700f und HST700r verwendet. Agarose-Gelelektrophorese, Gel gefärbt mit Ethidiumbromid. Bahn 1: DNA-Größenstandard. Bahn 2: frei.

Tabelle 4.3: Auflistung der Pflanzen, aus denen vermeintliche HST-Teilstücke gewonnen werden konnten. In den beiden letzten Spalten ist die Übereinstimmung der gefundenen Sequenzen zur HST-Referenzsequenz aus *Glechoma hederacea* (Sander, 2010) in Prozent angegeben.

| Familie      | Unterfamilie      | Art                                   | Übereinstimmung<br>DNA-Ebene | Übereinstimmung<br>Proteinebene |
|--------------|-------------------|---------------------------------------|------------------------------|---------------------------------|
| Blechnaceae  |                   | <i>Blechnum brasiliense</i>           | 86%                          | 90 %                            |
| Boraginaceae | Boraginoideae     | <i>Alkanna tuberculata</i>            | 100%                         | 100 %                           |
|              | Boraginoideae     | <i>Heliotropium arborescens</i>       | 56 %                         | 99 %                            |
|              | Boraginoideae     | <i>Lithospermum officinale</i>        | 100 %                        | 100 %                           |
|              | Boraginoideae     | <i>Lithospermum purpureocaeruleum</i> | 100 %                        | 100 %                           |
|              | Boraginoideae     | <i>Moltkia petraea</i>                | 85 %                         | 90 %                            |
|              | Boraginoideae     | <i>Myosotis arvensis</i>              | 100 %                        | 100 %                           |
|              | Boraginoideae     | <i>Onosma angustissima</i>            | 99 %                         | 98 %                            |
|              | Boraginoideae     | <i>Onosma albo-rosea</i>              | 56 %                         | 99 %                            |
|              | Boraginoideae     | <i>Symphytum officinale</i>           | 75 %                         | 72 %                            |
|              | Hydrophyllloideae | <i>Phacelia tanacetifolia</i>         | 85 %                         | 90 %                            |
| Lamiaceae    | Nepetoideae       | <i>Coleus blumei</i>                  | 84 %                         | 90 %                            |
|              | Nepetoideae       | <i>Dracocephalum moldavicum</i>       | 85 %                         | 91 %                            |
|              | Nepetoideae       | <i>Hyssopus officinalis</i>           | 100 %                        | 100 %                           |
|              | Nepetoideae       | <i>Lavandula angustifolia</i>         | 86 %                         | 89 %                            |
|              | Nepetoideae       | <i>Melissa officinalis</i>            | 96 %                         | 71 %                            |
|              | Nepetoideae       | <i>Nepeta cilicica</i>                | 85 %                         | 89 %                            |
|              | Nepetoideae       | <i>Nepeta nervosa</i>                 | 85 %                         | 91 %                            |
|              | Nepetoideae       | <i>Satureja montana</i>               | 84 %                         | 90 %                            |
|              | Lamioideae        | <i>Stachys cretica</i>                | 85 %                         | 90 %                            |
|              | Lamioideae        | <i>Stachys lavandulifolia</i>         | 85 %                         | 90 %                            |

#### **4.2.1. Alignment der Aminosäuresequenzen der amplifizierten HST-Teilstücke**

Die Analyse der aus den Nukleotidsequenzen translatierten Aminosäuresequenzen zeigt deutlich stark konservierte Bereiche, vor allem am Anfang und am Ende der putativen HST-Teilsequenzen (Abb. 4.8).

Die gefundenen Teilsequenzen (ca. 210 AS) liegen im Bereich des stark konservierten Motivs HxxxDG. Der verwendete 5'-Primer enthält diese konservierte Sequenz. Im Gegensatz zu den 3'-Primern zur Gewinnung von RAS-Sequenzen, enthält der 3'-Primer nicht die Sequenz DFGWG, sondern DFGGX. Nach dem Alignment und der Betrachtung mehrerer bekannter HST-Sequenzen, wurden verschiedene Varianten der DFGWG Sequenz im Labor getestet. Dabei stellte sich die Sequenz DFGGX als beste Sequenz für 3'-Primer zur Gewinnung von putativen HST-Sequenzen dar.

In Abbildung 4.8 ist der gewählte Bereich, innerhalb der HST-Sequenz von *Glechoma hederacea*, anhand eines Beispiels (Sequenz von *Stachys cretica*) dargestellt.



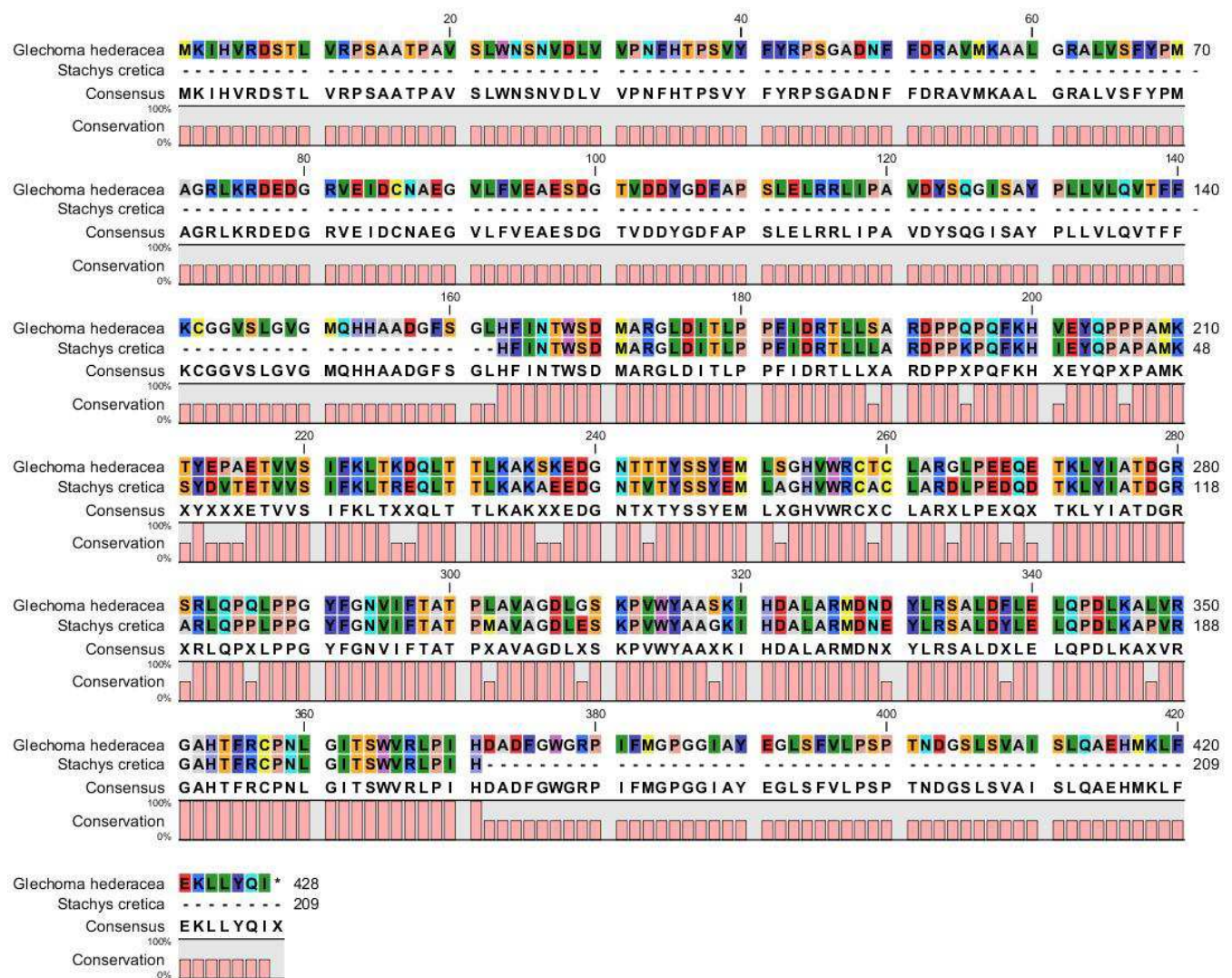


Abb. 4.8: Vergleich der AS-Gesamtsequenz einer HST aus *Glechoma hederacea* (Sander, 2010) mit den gefundenen Teilstücken, die zwischen den konservierten Sequenzen HxxxDG und DFGWG liegen. Beispielhaft wird hier die Teilsequenz von *Stachys cretica* gezeigt.

Die Sequenzlänge zwischen den konservierten Sequenzen HxxxDG und DFGWG bei der HST aus *Glechoma hederacea* als Referenz beträgt auch hier ungefähr 200 AS. Die gefundenen Teilstücke aus verschiedenen Pflanzenarten sind in Abbildung 4.9 gezeigt. Die Teilstücke besitzen, analog zu den RAS-Teilstücken, alle die ungefähr erwartete Länge von 200 AS. Stark konservierte Bereiche werden im Alignment nochmals hervorgehoben (Abbildung 4.9).

Hierbei sind die konservierten Bereiche am Anfang und Ende der Teilsequenz hervorzuheben. Weiterhin sind bei fast jedem Teilstück die Sequenzen TKLYIATDGR und VVSIFKLT (K/R) (D/E) Q (L/I) (außer *Symphytum officinale*) zu finden. Am Ende der Sequenz VVSIFKLT sind die basischen Aminosäuren Lysin und Arginin, sowie

die sauren Aminosäuren Asparaginsäure und Glutaminsäure zu finden. Gefolgt von Glutamin und den hydrophoben Aminosäuren Leucin oder Isoleucin. Generell findet man in den Teilsequenzen häufig den Austausch von Aminosäuren mit gleichen chemischen Eigenschaften, was im Alignment (Abbildung 4.9) an der identischen Hintergrundfarbe gut zu erkennen ist.

Die Primer werden im Alignment nicht mit angegeben. Das Alignment startet direkt nach dem 5'-Primer und endet vor dem 3'-Primer.



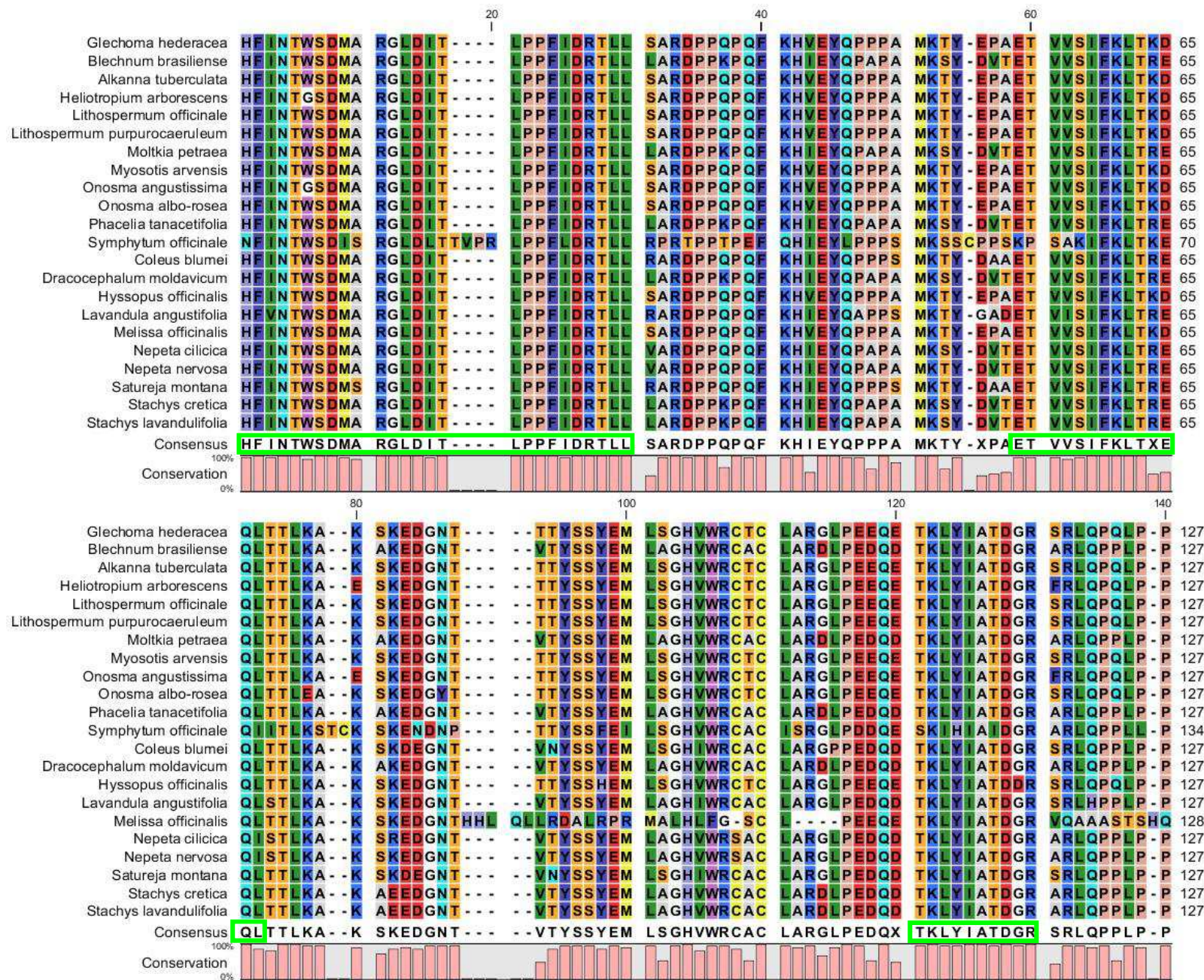






Abb. 4.9: Aminosäurealignment der amplifizierten Teilstücke einer putativen HST aus verschiedenen Arten der Lamiaceae und Boraginaceae (Pflanzen nach Verwandtschaftsverhältnissen geordnet). Als Referenz wird in der ersten Zeile die HST aus *Glechoma hederacea* gezeigt (Sander, 2010). Die am stärksten konservierten Abschnitte sind mit einem grünen Rahmen in der untersten Zeile „Consensus“ markiert.

#### **4.2.1. Alignment der DNA-Sequenzen der amplifizierten HST-Teilstücke**

Durch die zur Gewinnung von HST-Teilstücken optimierten Primer HST700f (forward) und HST700r (reverse) konnten aus zuvor hergestellten cDNA und gDNA-Präparationen verschiedener Pflanzenarten (Tabelle 4.3) ca. 700 bp große PCR-Amplifikate isoliert werden (Abbildung 4.10). Die Längenvariabilität der Teilstücke beträgt 3,0% (entspricht 21 bp).

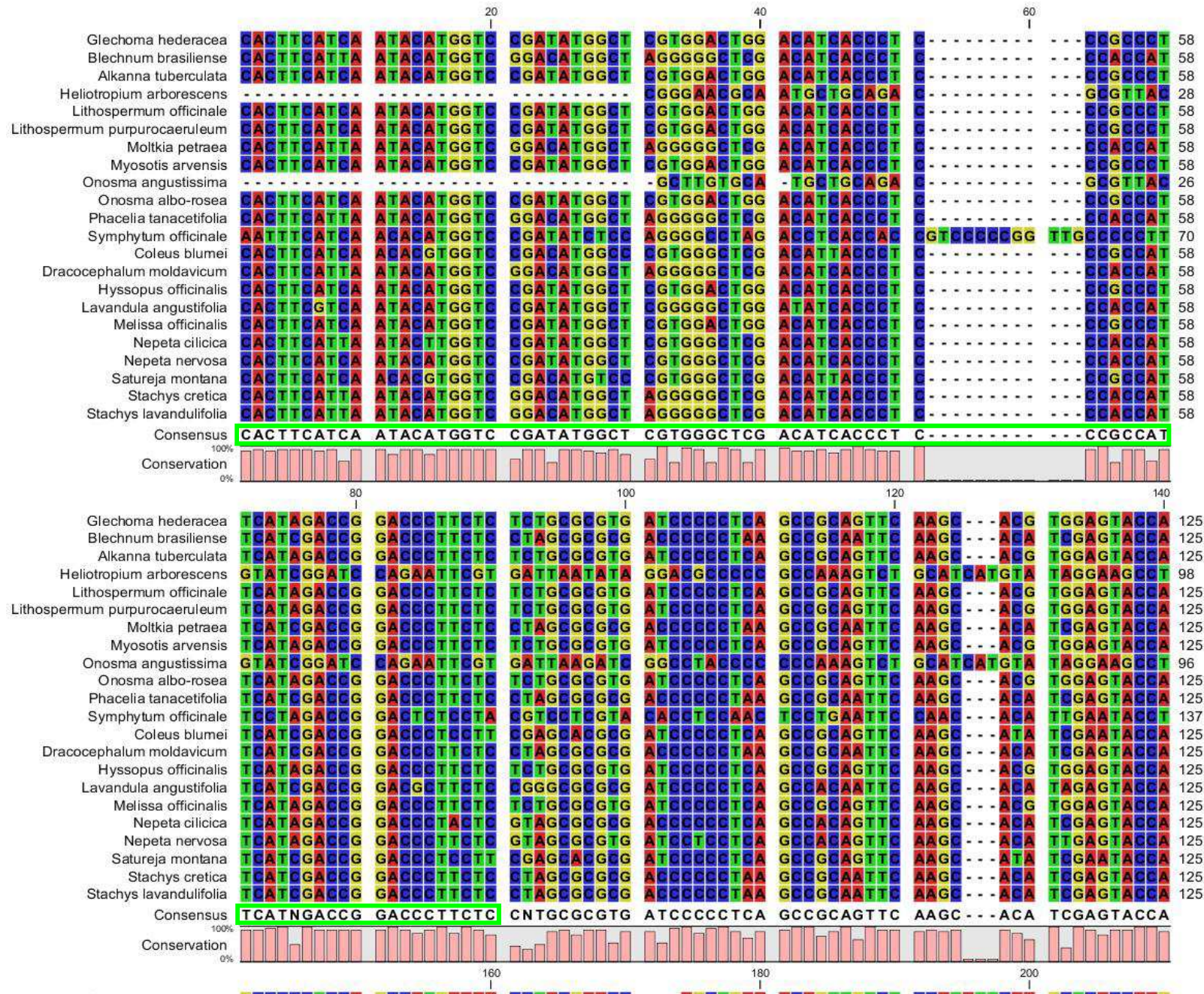
Der  $\chi^2$ -Test (Tabelle 4.4) zeigt die Homogenität in der Basenzusammensetzung der amplifizierten HST-Teilstücke ( $p = 1,00$ ).

Tabelle 4.4: Basenzusammensetzung der DNA-Sequenzen der amplifizierten HST-Teilstücke aus verschiedenen Pflanzen, nach Familienzugehörigkeit geordnet (siehe Tabelle 4.3). In der Spalte Länge sind die einzelnen Sequenzlängen des kompletten Alignments angegeben. Insgesamt weist der Datensatz eine deutlich homogene Basenzusammensetzung auf (p value = 1,00). Die statistische Auswertung des Datensatzes erfolgte mit Microsoft Excel.

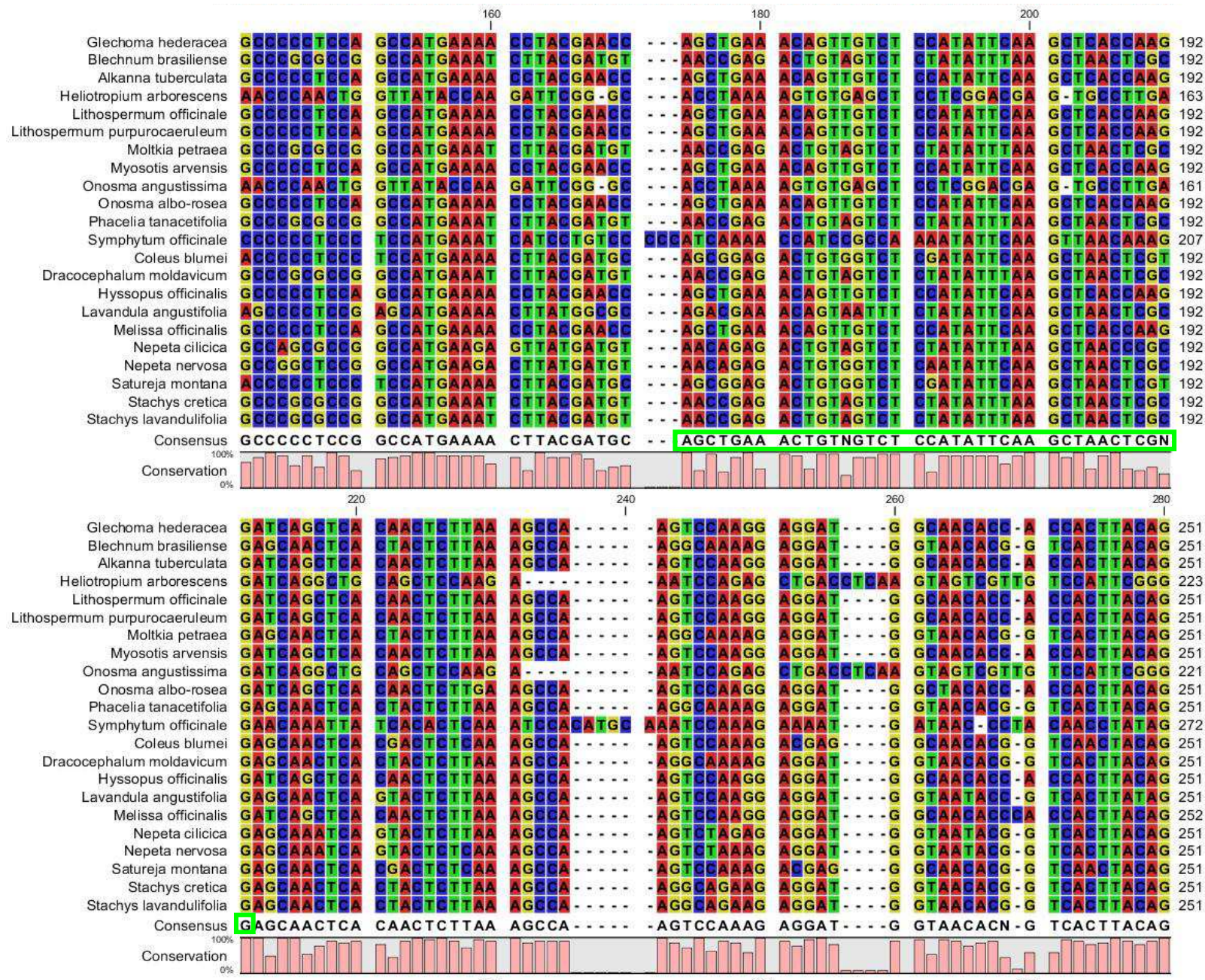
| <b>Pflanze</b>                        | <b>Adenin (A)</b> | <b>Cytosin (C)</b> | <b>Guanin (G)</b> | <b>Thymin (T)</b> | <b>Länge</b> |
|---------------------------------------|-------------------|--------------------|-------------------|-------------------|--------------|
| <i>Blechnum brasiliense</i>           | 0,224             | 0,288              | 0,265             | 0,223             | 627          |
| <i>Alkanna tuberculata</i>            | 0,243             | 0,281              | 0,241             | 0,235             | 627          |
| <i>Heliotropium arborescens</i>       | 0,235             | 0,249              | 0,274             | 0,242             | 616          |
| <i>Lithospermum officinale</i>        | 0,246             | 0,285              | 0,240             | 0,230             | 627          |
| <i>Lithospermum purpureocaeruleum</i> | 0,232             | 0,307              | 0,236             | 0,226             | 627          |
| <i>Moltkia petraea</i>                | 0,222             | 0,286              | 0,269             | 0,223             | 627          |
| <i>Myosotis arvensis</i>              | 0,248             | 0,280              | 0,241             | 0,231             | 627          |
| <i>Onosma angustissima</i>            | 0,235             | 0,252              | 0,273             | 0,240             | 614          |
| <i>Onosma albo-rosea</i>              | 0,245             | 0,283              | 0,240             | 0,232             | 627          |
| <i>Phacelia tanacetifolia</i>         | 0,221             | 0,287              | 0,266             | 0,227             | 627          |
| <i>Symphytum officinale</i>           | 0,270             | 0,226              | 0,233             | 0,271             | 645          |
| <i>Coleus blumei</i>                  | 0,205             | 0,330              | 0,266             | 0,199             | 627          |
| <i>Dracocephalum moldavicum</i>       | 0,234             | 0,277              | 0,255             | 0,235             | 627          |
| <i>Hyssopus officinalis</i>           | 0,234             | 0,302              | 0,241             | 0,223             | 627          |
| <i>Lavandula angustifolia</i>         | 0,231             | 0,282              | 0,263             | 0,224             | 627          |
| <i>Melissa officinalis</i>            | 0,219             | 0,285              | 0,255             | 0,240             | 611          |
| <i>Nepeta cilica</i>                  | 0,231             | 0,279              | 0,265             | 0,225             | 627          |
| <i>Nepeta nervosa</i>                 | 0,235             | 0,272              | 0,262             | 0,231             | 627          |
| <i>Satureja montana</i>               | 0,205             | 0,341              | 0,272             | 0,182             | 627          |
| <i>Stachys cretica</i>                | 0,243             | 0,277              | 0,247             | 0,232             | 627          |
| <i>Stachys lavandulifolia</i>         | 0,243             | 0,277              | 0,247             | 0,232             | 627          |
| Mittelwert                            | 0,233             | 0,283              | 0,255             | 0,229             | 625          |

Zur weiteren Übersichtlichkeit sind in dem folgenden DNA-Alignment die gleichen Teilbereiche markiert wie in dem Proteinalignment aus Abbildung 4.9.

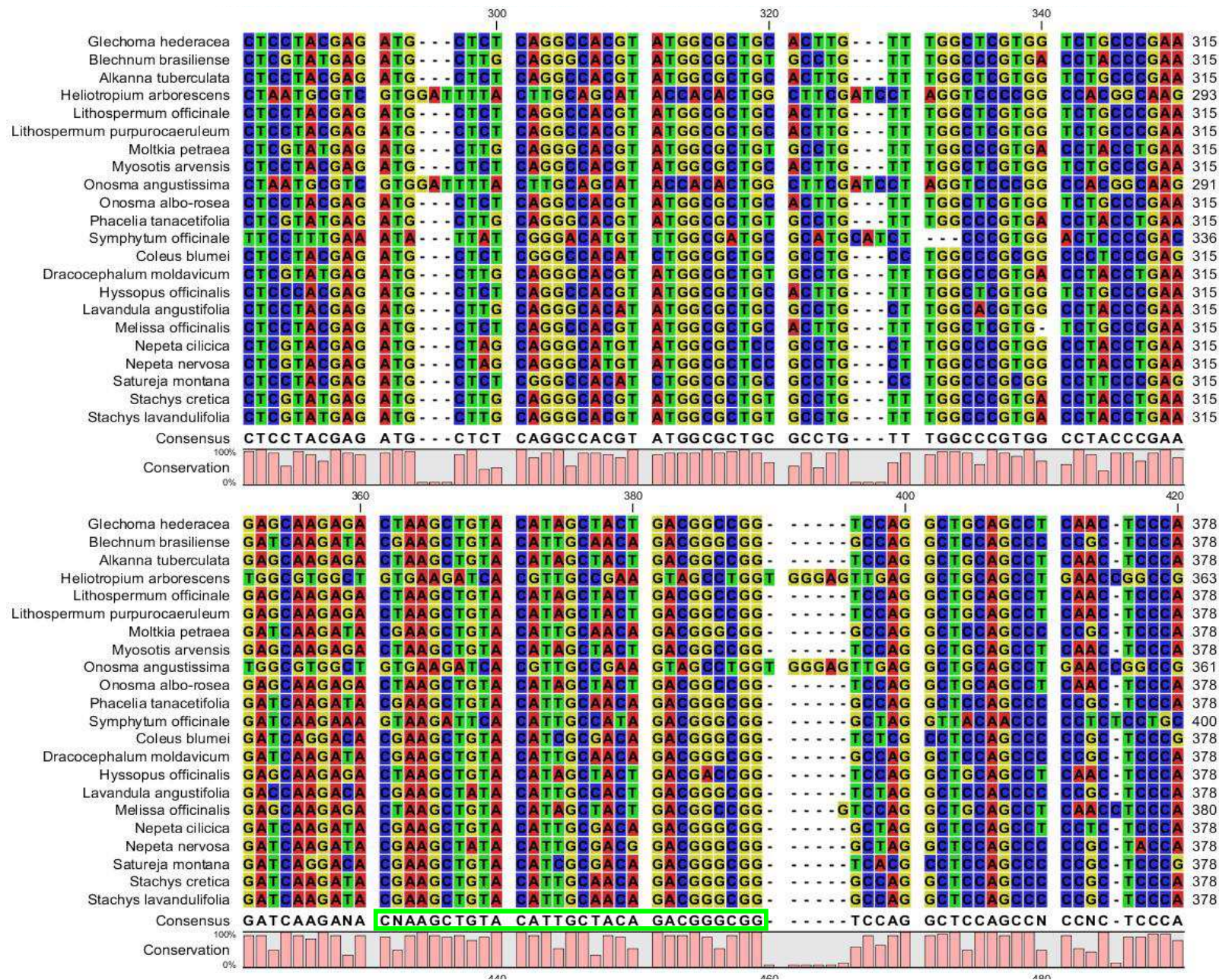








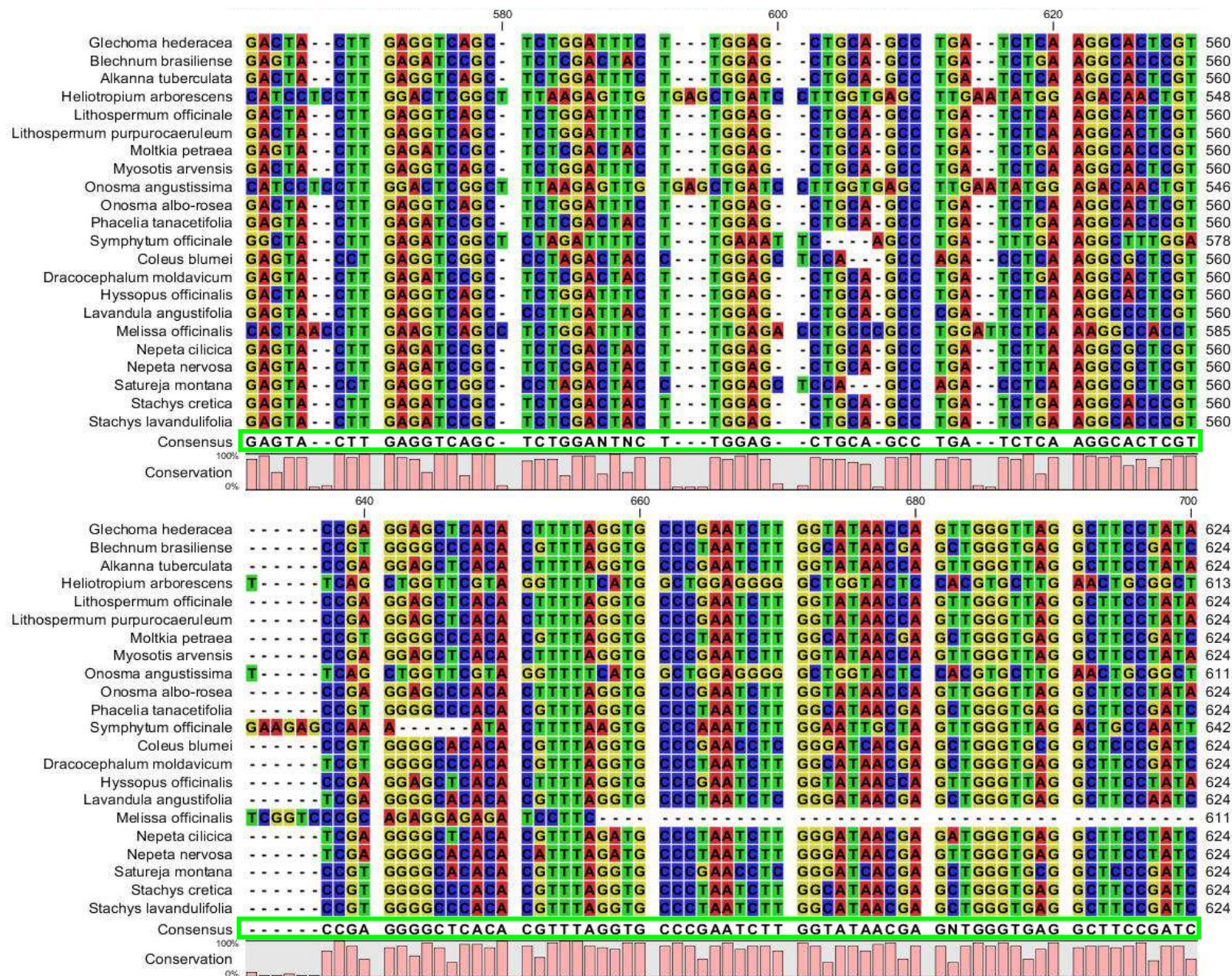












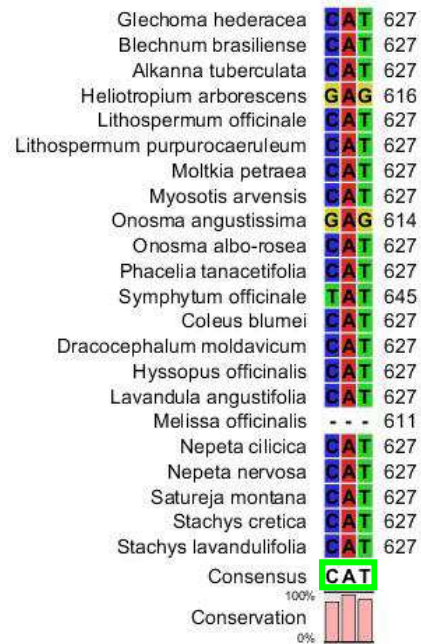


Abb. 4.10: DNA-Sequenzalignment der amplifizierten Teilstücke einer putativen HST aus verschiedenen Arten der Lamiaceae und Boraginaceae (Pflanzen nach Verwandtschaftsverhältnissen geordnet). Als Referenz wird in der ersten Zeile die HST aus *Glechoma hederacea* gezeigt (Sander, 2010). Die am stärksten konservierten Abschnitte sind mit einem grünen Rahmen in der untersten Zeile „Consensus“ markiert und entsprechen den markierten Bereichen des Proteinalignments aus Abbildung 4.9.



#### 4.3. Isolierung einer HST-ähnlichen Gesamtsequenz aus *Symphytum officinale*

Für Beinwell (*Symphytum officinale*, Boraginaceae) konnte eine putative HST-Gesamtsequenz amplifiziert werden. Zur Amplifikation wurden folgende Primer verwendet:

##### VL5-SO forward:

5'-TTA CAT ATG ATG AAG ATA GAT GTG CAA-3'

Translation: LHMMKIDVQ

##### VL3-SO reverse:

5'-GGA TCC TCA TCA TAT ATC ATA AAA GAT CTT CTG-3'

Translation: QKIFYDI\*\*GS

Die genomische Sequenz einer putativen HST-Gesamtsequenz aus *Symphytum officinale* konnte durch eine PCR mit den erwähnten Primern VL5-SO und VL3-SO, sowie einem zuvor frisch präparierten gDNA-Template isoliert werden (Abbildung 4.11). Die gDNA wurde aus Pflanzengewebe von *Symphytum officinale* gewonnen (siehe 3.1.1. Isolierung von genomischer DNA aus Pflanzenmaterial) und im Agarosegel nachgewiesen.

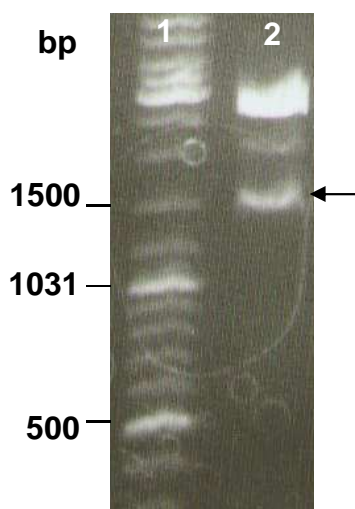


Abb. 4.11: Nachweis der ca. 1500 bp großen Gesamtsequenz einer putativen HST aus *Symphytum officinale* nach Ligation in pGEM®-T Vektor (Bahn 2), Transformation in *E. coli* der Linie DH5α und Verdau der Plasmidpräparation mit *NcoI* und *PstI*. Bahn 1: DNA-Größenstandard. Vektor von Bahn 2 enthält das gesuchte Insert und wurde sequenziert.

Die Analyse der erhaltenen Aminosäuresequenzen zeigt eine deutliche Ähnlichkeit (67%) zur Referenzsequenz aus *Glechoma hederacea*. Auch die stark konservierten Motive HxxxDG und DFGWG sind enthalten (Abb. 4.12). Werden Aminosäuren mit ähnlichen chemischen Eigenschaften (basisch, hydrophob, aromatisch etc.) in die Berechnung mit einbezogen ergibt sich eine weitaus höhere Ähnlichkeit von 82%.

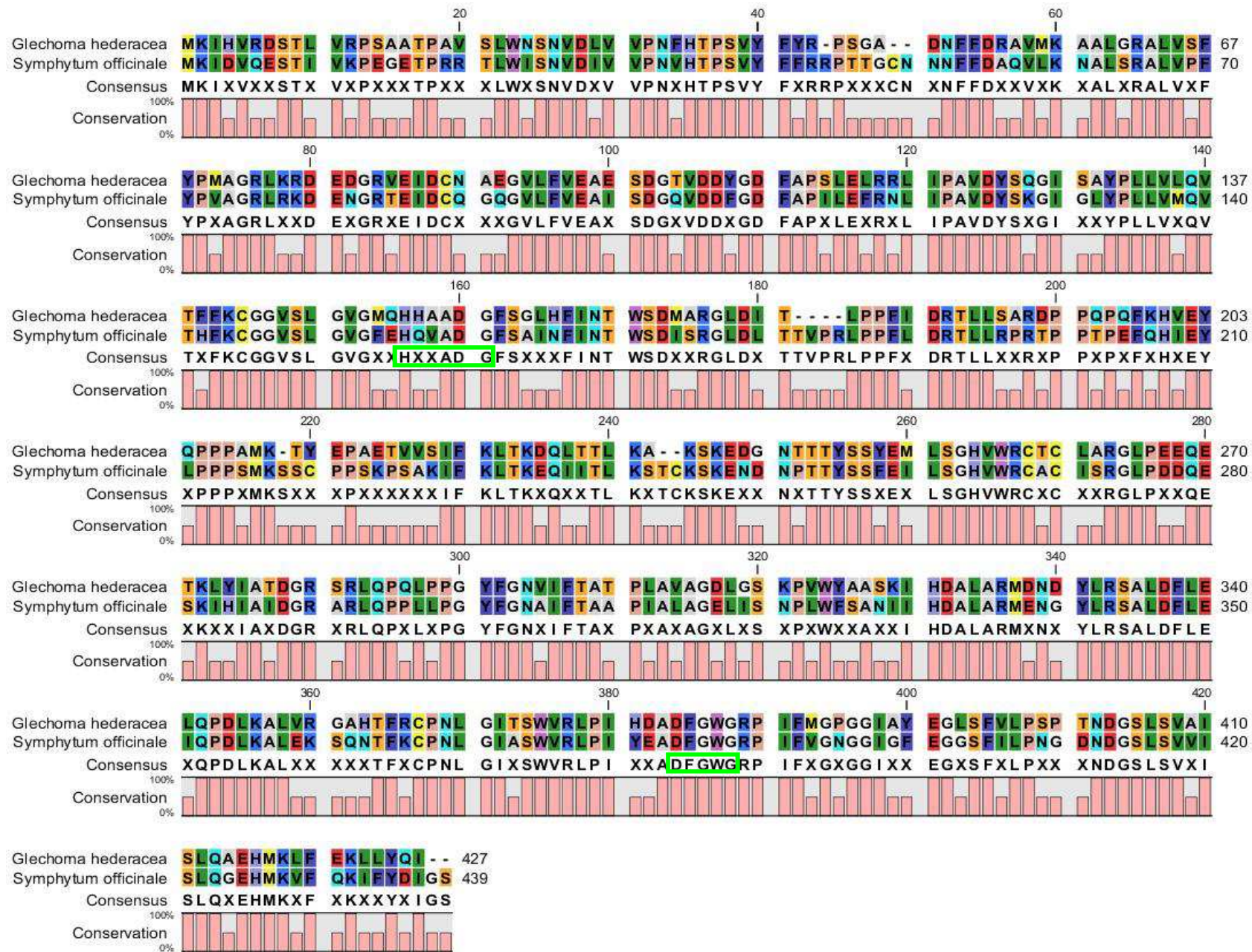


Abb. 4.12: Alignment der Vollängensequenz einer putativen HST-Proteinsequenz aus *Symphytum officinale* mit der

HST-Referenzsequenz aus *Glechoma hederacea*. Die konservierten Motive HxxxDG und DFGWG sind in der untersten Zeile „Consensus“ grün eingerahmt

Weiterhin zeigt die DNA-Sequenz der putativen HST aus *Symphytum officinale* auch eine gute Übereinstimmung (66%) mit der Referenzsequenz aus *Glechoma hederacea* (Abbildung 4.13).

Diese Daten sind ein Hinweis, dass die gefundene Sequenz aus *Symphytum officinale* ein HST-Enzym sein könnte.

Die genomische DNA (gDNA) kann in kodierende (Exons) und nichtkodierende (Introns) Sequenzen eingeteilt werden. Introns sind u.a. an der Steuerung der Replikation und der Transkription beteiligt, während Exons in Proteinstrukturen translatiert werden. Introns werden in eukaryotischen Zellen ebenfalls transkribiert, aber dann im Zellkern aus der prä-mRNA herausgespleißt, bevor die mRNA im Cytoplasma translatiert wird. Durch differentielles Spleißen kann ein Gen für verschiedene Proteinvarianten kodieren.

Die Untersuchung der Genstruktur von BAHD-Acyltransferasen zeigt, dass viele Vertreter ein typisches Q-Intron aufweisen. Dabei ist das Intron nach dem konservierten Glutamin (Q) 17 Aminosäuren vor dem HxxxD-Motiv eingefügt (St. Pierre und De Luca 2000).

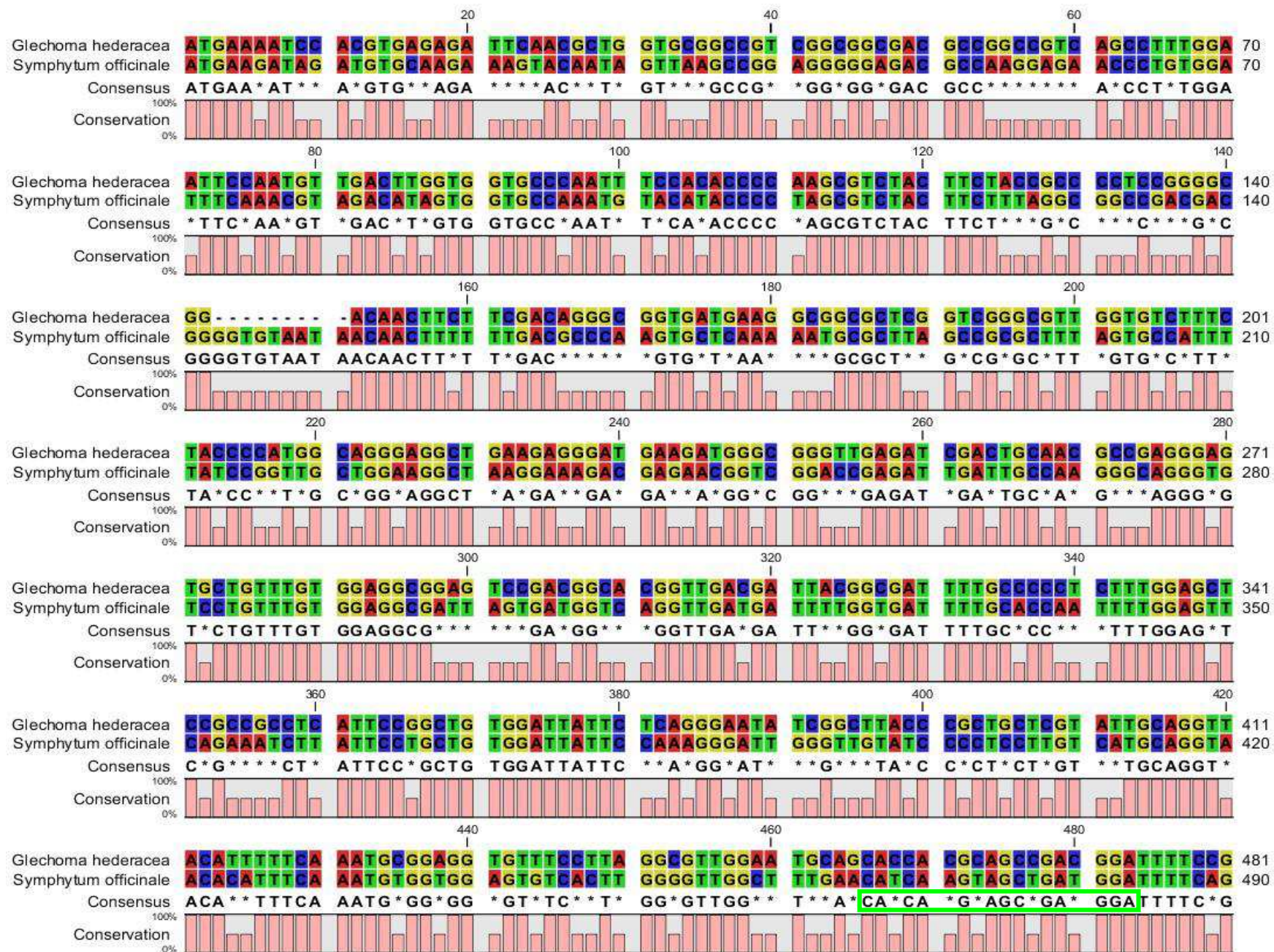
Dass trotz enormer Sequenzunterschiede innerhalb der BAHD-Acyltransferasen das Q-Intron in 75% aller Intron enthaltenden BAHD-Gene auftritt, gilt als Zeichen für einen gemeinsamen Vorfahren aller pflanzlicher BAHD-Acyltransferasen.

Mögliche Introns der isolierten DNA-Sequenz aus *Symphytum officinale*, wurden mit dem NetPlantGene Server berechnet (Hebsgaard et al., 2006). Die berechneten Introns werden mit Signifikanzwerten versehen. Signifikanzwerte von < 0.9 werden als sehr unwahrscheinlich angegeben. Bei der Sequenz aus *Symphytum officinale* wird ein Spleißdonor ab bp 947 angegeben (Signifikanzwert: 0,89). Die Sequenz im Bereich von bp 947 liegt dabei exakt im Bereich des typischen Q Introns. Der einzige mögliche Spleißakzeptor bei bp 979 besitzt auch einen niedrigen Signifikanzwert von 0,87. Da die Signifikanzwerte < 0.9 sind, wird angenommen, dass sich kein Intron oder Q-Intron innerhalb der Sequenz befindet.

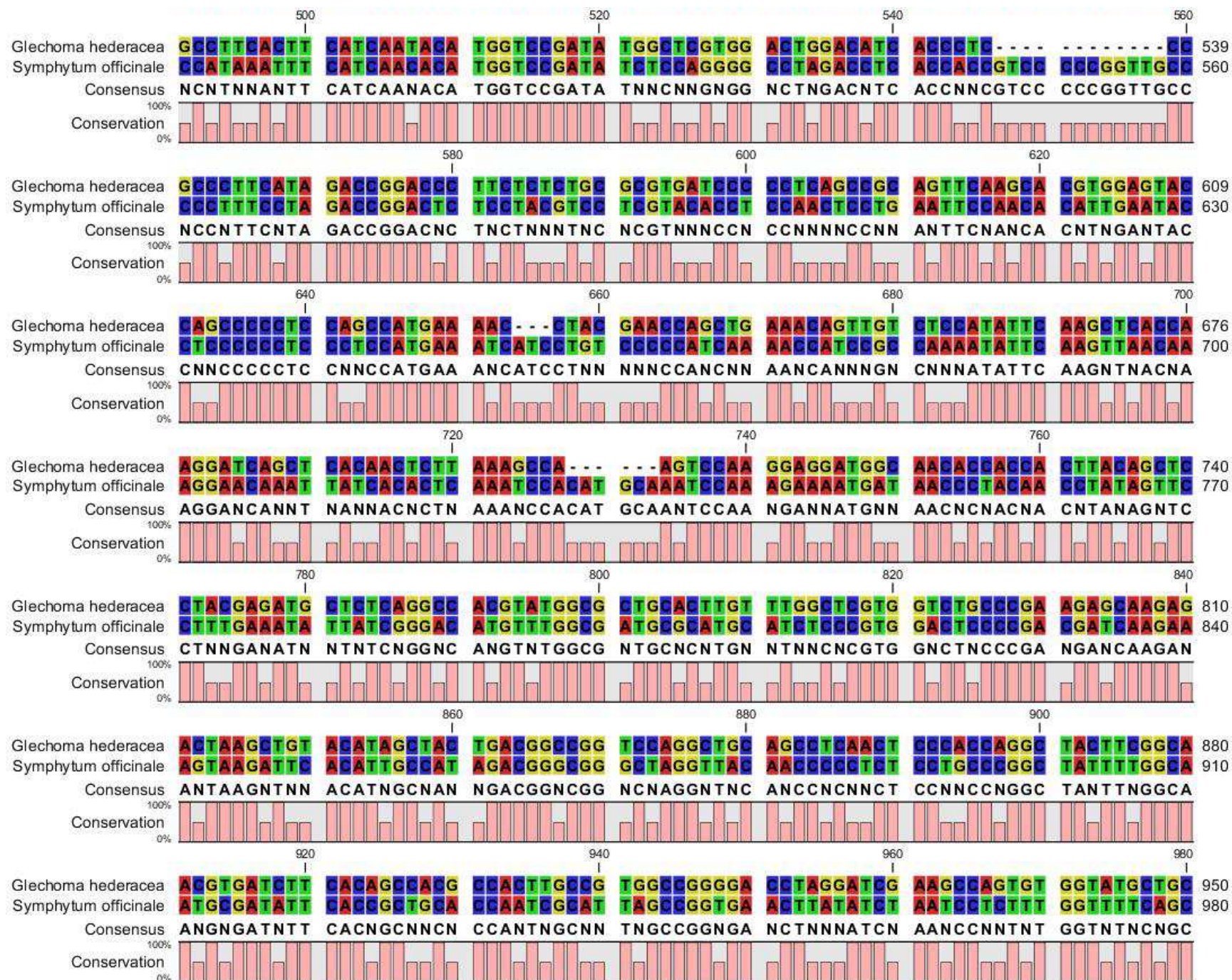
Ein weiterer Hinweis der nahen Verwandtschaft der beiden Proteine aus *Symphytum officinale* und *Glechoma hederacea*, ist die mathematische Modellierung der Aminosäuresequenz zur Bestimmung der dreidimensionalen Tertiärstruktur. Hier zeigt sich eine gute Übereinstimmung sowohl der Sekundär-, als auch der Tertiärstruktur der Proteine. Beide Proteine enthalten 9 Helices und 6 parallel zueinander angeordnete  $\beta$ -Faltblatt Strukturen. Die konservierte Sequenz HxxxDG bildet bei beiden Proteinen im inneren Teil ihrer Struktur eine Schleifenform aus. Die konservierte Sequenz DFGWG liegt in einem weiter nach außen gefassten Bereich in einer Bogenform, relativ exponiert an der Oberfläche beider Proteine (Abbildung 4.14).

Die Expression einer aktiven Form des Enzyms konnte jedoch nicht erzielt werden. Expressionsversuche und HST-Aktivitätstests, sowie der folgende Nachweis über HPLC-Analyse waren leider nicht erfolgreich. In den Kapiteln 5.1 Methodik und 5.4 Gewinnung einer putativen HST-Volllängensequenz aus Beinwell (*Symphytum officinale*) werden die Probleme, die z.B. bei den Versuchen zur Herstellung von cDNA entstanden sind, beschrieben.











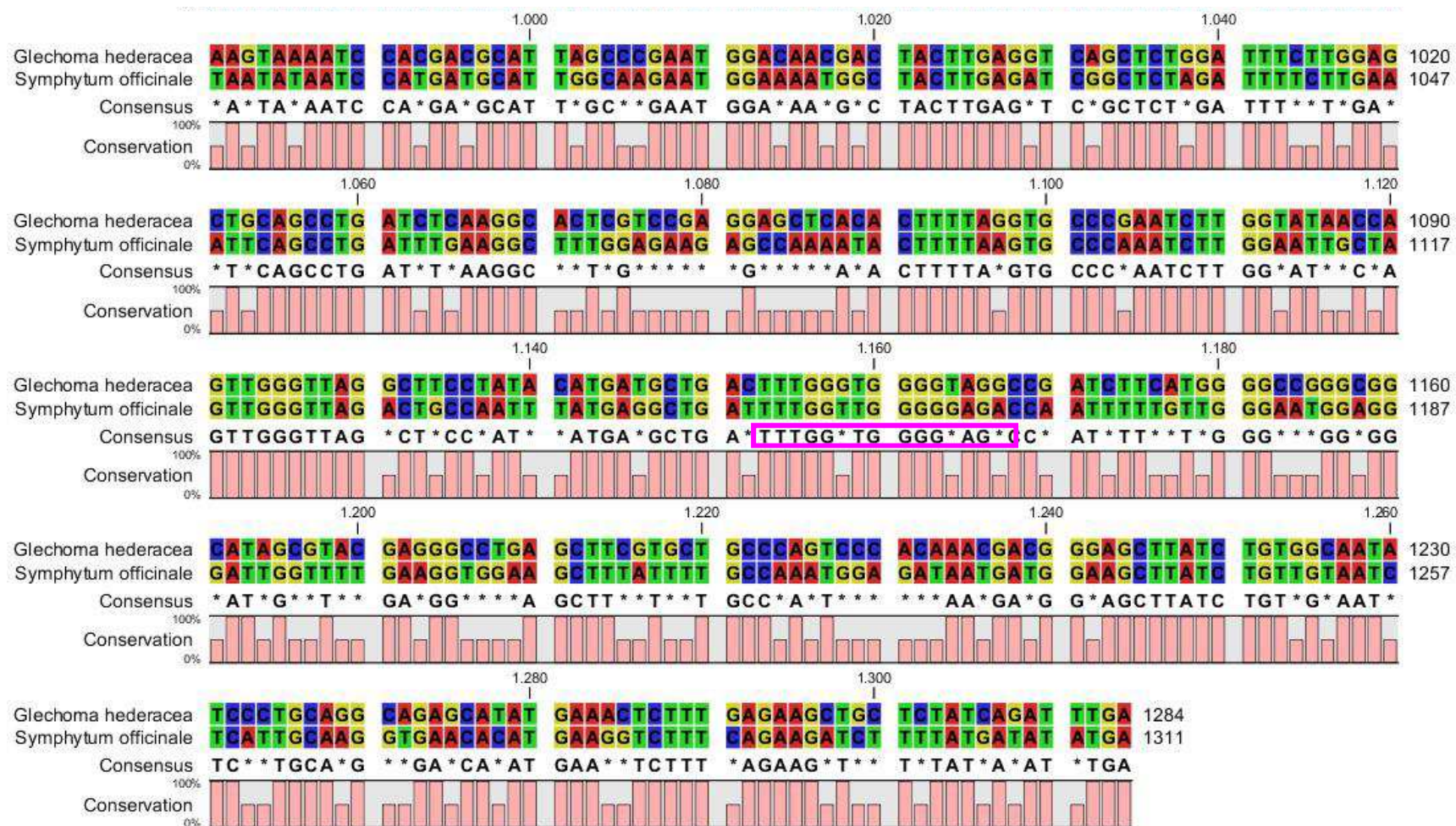


Abb. 4.13: Alignment der putativen HST-DNA-Sequenz aus *Symphytum officinale* mit der HST-Referenzsequenz aus *Glechoma hederacea*. Die konservierten Motive HxxxDG (grün) und DFGWG (violett) sind in der untersten Zeile „Consensus“ eingerahmt

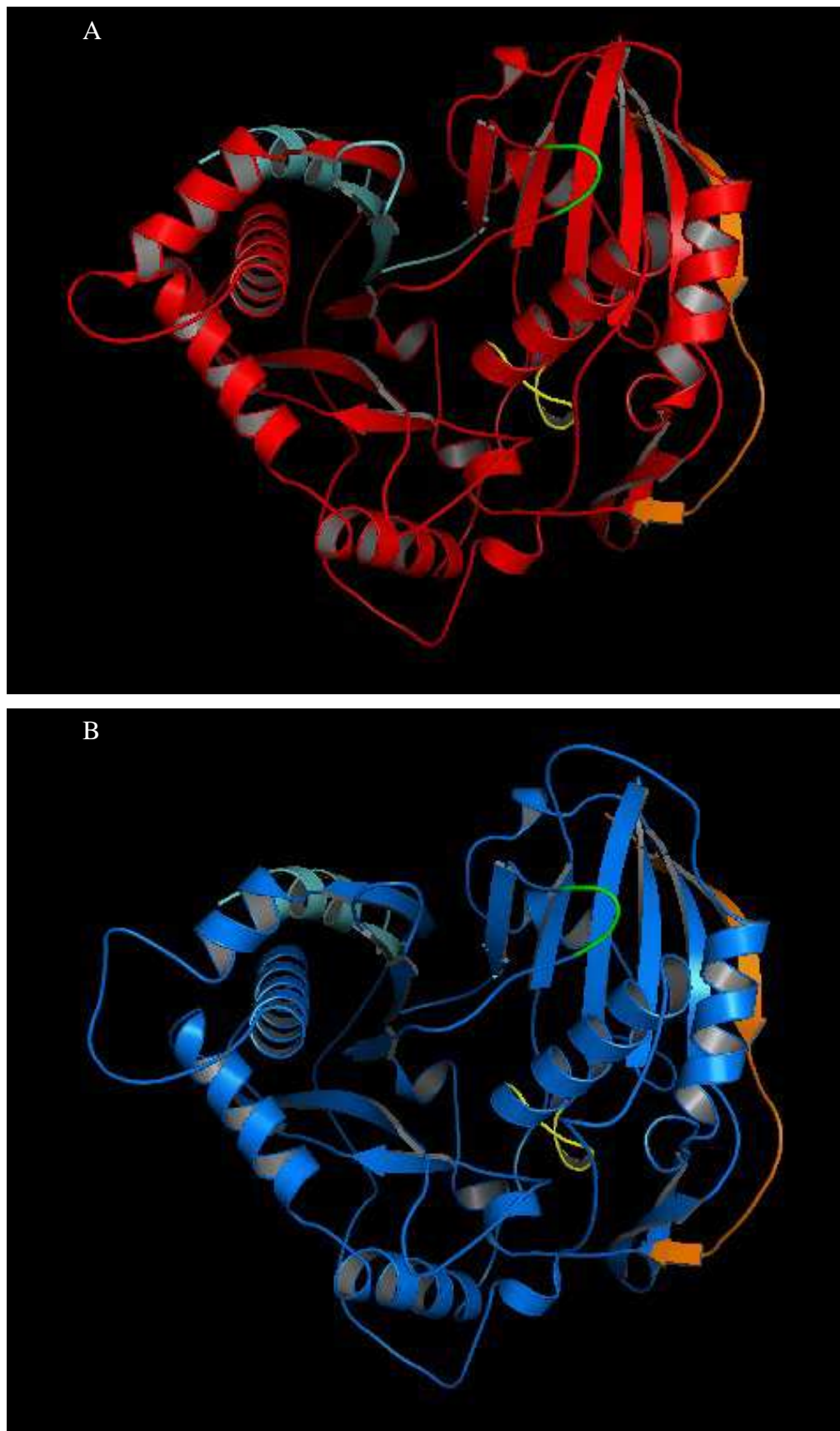


Abbildung 4.14: Berechnete Tertiärstruktur der HST aus *Glechoma hederacea* (A, rot) und der putativen HST aus *Symphytum officinale* (B, blau), mittels Swiss-Model ([swissmodel.expasy.org](http://swissmodel.expasy.org)). Der N-Terminus ist bei beiden Strukturen orange, und der C-

Terminus hellblau gefärbt. Das konservierte Motiv HxxxDG ist gelb gefärbt. Die konservierte Sequenz HxxxDG bildet bei beiden Proteinen im inneren Teil ihrer Struktur eine Schleifenform aus. Die konservierte Sequenz DFGWG ist grün gefärbt und befindet sich oberhalb der parallel angeordneten  $\beta$ -Faltblatt Strukturen, in einer Bogenform, relativ exponiert an der Oberfläche. Die Darstellung wurde mit der frei erhältlichen Education Software Pymol ([pymol.org](http://pymol.org)) erzeugt.

## 5. Diskussion

### 5.1. Methodik

In der vorliegenden Arbeit können die gewonnenen Ergebnisse in folgende Struktur gegliedert werden:

- Gewinnung von putativen RAS-Teilsequenzen aus verschiedenen, nahe verwandten Pflanzenarten
- Gewinnung von putativen HST-Teilsequenzen aus verschiedenen, nahe verwandten Pflanzenarten
- Gewinnung einer Volllängensequenz einer vermeintlichen Hydroxycinnamoyltransferase aus Beinwell (*Symphytum officinale*)
- Erstellung eines phylogenetischen Baums aus den gewonnenen RAS- und HST-Teilsequenzen
- Diskussion der Verwandtschaftsverhältnisse auf der Basis der isolierten RAS- und HST-Sequenzen

Für alle Versuche waren die Isolierung von mRNA und anschließende cDNA-Synthese, RACE-PCR oder die Isolierung von genomischer DNA notwendig.

Bei der Herstellung von mRNA nach Chomczynski und Sacchi (1987) zeigte sich, dass die im Labor angewandte Methode nicht immer befriedigende Ergebnisse erzielte. Häufig war nach Kontrolle der mRNA mittels Gelelektrophorese deren Integrität nicht gegeben, oder es konnten nicht alle Verunreinigungen eliminiert werden. Erst nach Bestätigung einer ausreichenden mRNA-Qualität ist jedoch der weiterführende Schritt der cDNA-Synthese sinnvoll. Sowohl bei der Herstellung von cDNAs, als auch bei der spezifischen Amplifikation von definierten cDNAs mittels RACE-PCR konnten nur unzureichende Amplifikate erzeugt werden. Nach Überprüfung sämtlicher Prozessschritte und Ausgangsmaterialien, war ein Hauptgrund die häufig nicht ausreichende Qualität der zuvor hergestellten mRNA.

Bei der weiteren Herstellung von cDNAs wird das Enzym Reverse Transkriptase (RT) verwendet. Für die praktische Durchführung wurde die thermostabile RT „Revert Aid™ M-MuLV“ verwendet. Dieses Enzym ermöglicht die Umschreibung von mRNA zu cDNA. Nach dieser Erststrangsynthese kann die gewonnene cDNA als Vorlage für

eine anschließende PCR verwendet werden. Vorteil der Verwendung von cDNA ist das Fehlen von Introns. Die verwendete mRNA für die Erststrangsynthese ist im Idealfall vollständig gespleißt. Jedoch kann es bei der Synthese von mRNA zu Abbrüchen oder Abbau kommen, so dass die RT unvollständig kodierende cDNA erzeugt. Eine bessere Qualität der cDNA hätte durch z.B. Größenfraktionierung mittels Größenausschlusschromatographie erzielt werden können. Dabei werden kurze Sequenzen retardiert und von der übrigen cDNA getrennt. Ein weiterer Vorteil von cDNAs ist deren Stabilität und Lagerfähigkeit. Durch die beschriebenen Probleme bei der Herstellung von geeigneter mRNA konnten jedoch auch bei der cDNA-Synthese häufig nur unzureichende Ergebnisse erzielt werden.

Als vorrangige Methode wurde somit die Isolierung von genomischer DNA (gDNA) nach Rogers und Bendich (1985) gewählt. Diese Methode ist im Allgemeinen einfacher umzusetzen und der Prozess unanfälliger gegenüber Störfaktoren. Ein Nachteil ist, dass die kodierenden Sequenzen der gDNA noch durch Introns unterbrochen sein können. Die Ausbeute und Reinheit wurden über spektrophotometrische Messung und Gelelektrophorese kontrolliert und zeigten meist gute Qualität. Obwohl bei der Isolierung von gDNA immer auch DNA aus Plastiden und Mitochondrien vorhanden sein kann und die Introns noch enthalten sind, konnten unter Einsatz von gDNA die meisten in dieser Arbeit vorgestellten Sequenzen gewonnen werden.

## **5.2. Gewinnung von putativen RAS-Teilsequenzen aus verschiedenen Pflanzenarten**

Die Rosmarinsäure Synthase (Hydroxycinnamoyl-CoA:Hydroxyphenyllactat Hydroxycinnamoyltransferase; rosmarinic acid synthase; RAS) ist ein essentielles Enzym zur Synthese von Rosmarinsäure.

Rosmarinsäure kommt in der Familie der Boraginaceae durchgehend vor. In der Familie der Lamiaceae ist Rosmarinsäure häufig zu finden, jedoch überwiegend in der Unterfamilie der Nepetoideae, hier jedoch in – soweit heute bekannt – allen Pflanzenarten. Verschiedene Studien belegen, dass die Rosmarinsäure ein charakteristischer Sekundärmetabolit der Nepetoideae ist (Litvinenko et al., 1975; Grayer et al., 2003). In weiteren Unterfamilien der Lamiaceae ist Rosmarinsäure

auch in einzelnen Arten zu finden (z.B. *Teucrium scorodonia*, *Aegiphila mollis* und *Hymenopyramis brachiata* (Petersen et al., 2009)), kann aber in keiner weiteren Unterfamilie als charakteristischer Sekundärmetabolit angesehen werden (Pedersen, 2000). Chemotaxonomisch ist die Rosmarinsäure für die Familie der Lamiaceae somit nur bedingt geeignet. Weiterhin wurde Rosmarinsäure in unterschiedlichsten Familien nachgewiesen, von den evolutionär frühen Landpflanzen, den Hornmoosen (Takeda et al., 1990; Petersen et al., 2003), bis hin zu höher entwickelten Pflanzenarten (Harborne, 1966; Aquino et al., 1990; Ly et al., 2006). Teilweise ist die Rosmarinsäure nur in wenigen Gattungen einer Familie zu finden und innerhalb einer Gattung oft auch nicht in allen Arten. Ein evolutionär einheitliches Bild der Synthese von Rosmarinsäure ist aus den momentan sehr gestreuten Daten nur schwer abzulesen.

Hier können molekularbiologische Untersuchungen der RAS einen weiteren Baustein bereitstellen. Von den insgesamt acht Enzymen zur Synthese von Rosmarinsäure sind maximal vier spezifisch für diesen Syntheseweg. Eines dieser Enzyme ist die RAS, welche erstmals 1988 von Petersen und Alfermann nachgewiesen und 2006 von Berger et al. kloniert wurde. Sie überträgt die CoA-aktivierten Säuren 4-Cumaroyl-CoA oder Caffeoyl-CoA von CoA auf 4-Hydroxyphenyllactat oder 3,4-Dihydroxyphenyllactat, nicht aber auf Shikimat oder Chinat (Petersen et al., 2009).

Es ist zu prüfen, ob in evolutionär weit auseinander liegenden Pflanzenarten RAS-Sequenzen zu finden sind. Sollte dies der Fall sein, könnte bei hohen Sequenzidentitäten der RAS-Sequenzen davon ausgegangen werden, dass die Fähigkeit, Rosmarinsäure zu produzieren eine evolutionär früh erworbene Eigenschaft ist. Viele Arten hätten dann jedoch im Laufe der Evolution diese Fähigkeit wieder verloren. Als zweite Möglichkeit könnte sich die Fähigkeit der Synthese von Rosmarinsäure auch mehrfach unabhängig entwickelt haben. Dies wäre der Fall, wenn nachgewiesene Enzyme mit RAS-Aktivität, kaum oder keine Sequenzähnlichkeit zu den bislang aus den Lamiaceae bekannten RAS-Genen aufweisen würden.

Von der Sequenz allein kann jedoch nicht unmittelbar auf die Substratspezifität des Enzyms geschlossen werden. Nur durch eine nachfolgende biochemische Charakterisierung kann gezeigt werden, ob es sich um eine RAS, HST oder HQT handelt oder ob gänzlich andere Substrate akzeptiert werden. Dafür muss die vollständige Sequenz des Enzyms amplifiziert und nachfolgend ein aktives Enzym



heterolog exprimiert werden. In dieser Arbeit konnten nur putative RAS-Teilstücke aus verschiedenen Pflanzen gewonnen werden. Eine Vollängensequenz konnte nicht isoliert werden.

Viele Hydroxycinnamoyltransferasen enthalten ein Q-Intron, welches nach dem konservierten Glutamin (Q) 17 Aminosäuren vor dem HxxxD-Motiv eingefügt ist (St. Pierre und De Luca, 2000). Die hier amplifizierten Teilstücke starten direkt hinter dem HxxxD-Motiv und sollten somit dieses Intron nicht enthalten. Die Längen der Teilstücke wurden mit ca. 700 bp angenommen, und sie konnten innerhalb dieses Größenbereichs auch amplifiziert werden. Die für diese Annahme notwendige RAS-Sequenz wurde zuvor in unserem Labor aus *Coleus blumei* ermittelt (Berger et al., 2006).

Die amplifizierten Teilstücke zeigten auf Aminosäureebene große Ähnlichkeiten zu der Referenzsequenz der RAS aus *Coleus blumei* (92% bis 100%). Die AS-Sequenz aus *Lamium album* ist als einzige Sequenz zu 100% identisch mit der Referenzsequenz aus *Coleus blumei*.

Eine von insgesamt zwei translatierten AS-Sequenzen aus *Ajuga reptans* zeigte als einzige Sequenz nur sehr geringe Ähnlichkeit zur Referenzsequenz (50%). Auf Ebene der Nukleotidsequenzen ist eine Ähnlichkeit von 81% bis 99% vorzufinden. Auch hier zeigt die Sequenz aus *Ajuga reptans* nur eine geringe Ähnlichkeit.

Weiterhin eröffnet die Analyse der erhaltenen Aminosäuresequenzen stark konservierte Bereiche innerhalb dieser Teilsequenz, die sich auch auf Ebene der Nukleotidsequenzen wiederfinden. Besonders hervorzuheben ist dabei die Sequenz HIPFDG (ab AS 279), welche in allen Sequenzen (außer *Ajuga reptans* 2) zu finden ist. Interessanterweise beinhaltet diese Sequenz die gleiche Aminosäurenreihenfolge wie das bereits bekannte stark konservierte Motiv HxxxDG. Zusätzlich sind noch drei weitere konservierte AS-Sequenzen (außer bei *Ajuga reptans* 1) zu finden (siehe Abb. 4.5):

|                     |                    |
|---------------------|--------------------|
| ab Position AS 251: | STFEVLAGHIWR       |
| ab Position AS 295: | FGNAIFFATP         |
| ab Position AS 352: | AQGAHSFRCPNLWVISWV |

(Die Nummerierung der AS bezieht sich auf die RAS-Referenzsequenz aus *Coleus blumei*).

Die größten Unterschiede innerhalb dieser drei konservierten Sequenzbereiche zeigen *Ajuga reptans* 1 und *Mentha X piperita*, wobei *Mentha X piperita* Rosmarinsäure enthält, *Ajuga reptans* jedoch nicht.

Bei der Gleichsetzung von Aminosäuren mit ähnlichen chemischen Eigenschaften (sauer, basisch, aromatisch etc.), stimmen die Identitäten der Sequenzen jedoch noch besser überein (siehe Tabelle 4.1).

Abbildung 5.1 zeigt die betreffenden Sequenzabschnitte innerhalb eines berechneten 3D-Modells der RAS-Referenzsequenz aus *Coleus blumei*. Es ist sehr wahrscheinlich, dass die gefundenen Sequenzen eine identische Struktur innerhalb des Moleküls aufweisen. In Kapitel 5.5 „Erstellung eines phylogenetischen Baums der verwendeten Pflanzen“ wird nochmals auf die identische Protein-Modellstruktur bereits bekannter unterschiedlicher BAHD-Acyltransferasen eingegangen werden. In Abbildung 5.1 ist ein Tryptophanrest (AS 370, blau) vor dem Eingang des Kanals zum Aktivitätszentrum zu erkennen. Dieser Tryptophanrest ist nicht stark konserviert, und man findet an dieser Position entweder Tryptophan- oder Argininreste bei allen amplifizierten RAS oder HST Teilsequenzen (siehe Abbildung 4.5). Für HST Enzyme ist bekannt, dass das Arginin (sowie drei weitere Aminosäurereste) an dieser Position die Phosphatgruppen des Coenzym-A-Liganden binden (Lallemant et al., 2012b). Weiter ist zu erkennen, dass die am stärksten konservierten Sequenzen dem Kanal des aktiven Zentrums zuzuordnen sind. Dies deutet darauf hin, dass die Form und Ausprägung des Kanals wichtig für die spezifische Funktion der RAS ist. Diese Vermutung wird im nächsten Kapitel weiter ausgeführt.

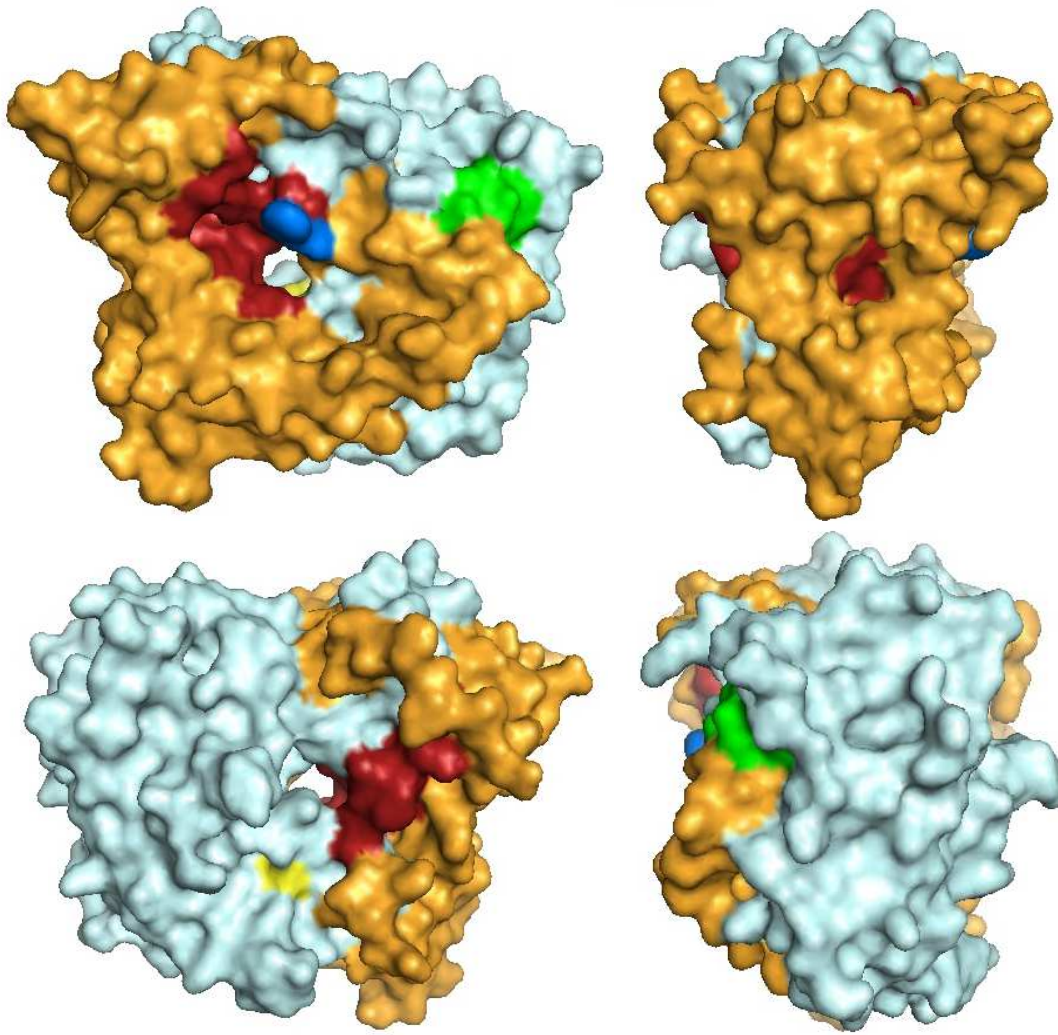


Abb. 5.1: Oberflächendarstellung der RAS aus *Coleus blumei* mit eingefärbten Sequenzabschnitten. Das Molekül wird in Leserichtung pro Bild um 90° rechtsseitig gedreht. Weiß: Nicht amplifizierter Sequenzabschnitt aus der Volllängeensequenz von *Coleus blumei*, orange: amplifizierter Sequenzabschnitt, rot: stark konservierte Sequenzbereiche, gelb: katalytisches Zentrum mit dem Motiv HxxxDG, grün: charakteristisches Motiv DFGWG, blau: Tryptophanrest, vor dem Kanal zum katalytischen Zentrum angeordnet. Die am stärksten konservierten Sequenzen (rot) sind dem Kanal des aktiven Zentrums zuzuordnen. Die Darstellung wurde mit der frei erhältlichen Education Software Pymol ([pymol.org](http://pymol.org)) erzeugt.

### **5.3. Gewinnung von putativen HST-Teilsequenzen aus verschiedenen Pflanzenarten**

Die Hydroxycinnamoyl-CoA:Shikimat Hydroxycinnamoyltransferase (HST) ist ein essentielles Enzym zur Synthese von Caffeoyleshikimat (Schoch et al., 2001; Hoffmann et al., 2003), welches ein Intermediat im Biosyntheseweg von Monolignolen ist (Hoffmann et al., 2004, Hoffmann et al., 2005). Die Fähigkeit, Lignin zu synthetisieren, ist unter fast allen landlebenden Pflanzen (außer Moosen) verbreitet. Daher muss sich die Ligninbiosynthese evolutionär früh entwickelt haben. Martone et al. (2009) identifizierten Lignin sogar als Zellwandbestandteil in frühevolutionären Rotalgen der Art *Calliarthron cheliosporioides*. Diese Eigenschaft allerdings soll, nach heutigem Kenntnisstand, evolutionär unabhängig entstanden sein. Die Ligninmonomere (Monolignole) teilen sich auf in 4-Cumarylalkohol (H-Einheit), Coniferylalkohol (G-Einheit) und Sinapylalkohol (S-Einheit). Die Zusammensetzung der Lignine unterscheidet sich bei den Pflanzenarten: die prozentualen Anteile der verschiedenen Monolignole und der Vernetzungsgrad sind sehr verschieden (Boudet, 2000).

Hoffmann et al. (2004) zeigten durch RNAi-basierte Blockierung der HST in *Nicotiana benthamiana* und *Arabidopsis thaliana* die Funktion der HST bei der Ligninbiosynthese. Die Ligninbiosynthese war bei beiden Pflanzenarten durch die Blockierung der HST gehemmt. Das Verhältnis von 4-Hydroxyphenyl-Lignin zu Syringyl-Lignin verschiebt sich im Vergleich zu den Wildtyp-Pflanzen zugunsten von 4-Hydroxyphenyl-Lignin. Die HST ist an der Biosynthese des höher methoxylierten Syringyl-Lignins beteiligt, während 4-Hydroxyphenyl-Lignin direkt aus 4-Cumarylalkohol gebildet wird. Diese Versuche unterstützen den momentan postulierten Monolignolbiosyntheseweg. Die Beteiligung der HST an der Produktion der höher substituierten Monolignole in *Pinus radiata* (Wagner et al., 2007) und *Medicago sativa* (Shadle et al., 2007) konnte ebenfalls gezeigt werden. Pflanzenarten, die kein Syringyl-Lignin produzieren, ändern jedoch ebenfalls das Substitutionsmuster des Lignins nach Blockieren der HST. Als Modellpflanze für diesen Versuch diente die Gymnosperme *Pinus radiata*, welche kein Syringyl-Lignin produziert. Wurde die Funktion der HST durch RNA-Interferenz ausgeschaltet, kam es zu signifikanten Verschiebungen in Gehalt und Zusammensetzung des Lignins. Es wurde eine Erhöhung der 4-Hydroxyphenyl- und eine Reduktion der Guaiacyl-

Einheiten beobachtet (Wagner et al., 2007). Im Weiteren kann Caffeoyl-CoA in *meta*-Position methyliert werden (CCoAOMT: Caffeoyl-CoA O-Methyltransferase), so dass mit Feruloyl-CoA (Fer-CoA) die Vorstufe von Coniferylalkohol (G-Einheit) und Sinapylalkohol (S-Einheit) vorhanden ist (Hoffmann et al., 2004).

Für die HST konnte zusätzlich auch die Katalyse der Rückreaktion nachgewiesen werden: Sie setzt Caffeoyl-CoA und Shikimat in Anwesenheit von CoA aus Caffeoylshikimat frei.

Im Gegensatz dazu zeigten Versuche mit Tomaten, in denen die Expression der nahe verwandten Hydroxycinnamoyl-CoA:Chinat Hydroxycinnamoyltransferase (HQT) unterdrückt wird, dass die Ligninproduktion hierdurch nicht beeinflusst wird (Niggeweg et al., 2004). Dieser Versuch zeigt u.a., dass von einer generellen Beteiligung der HQTs an der Ligninsynthese nicht auszugehen ist. Die essentielle Funktion der HQT in der Synthese von Chlorogensäurederivaten ist dagegen sehr gut zu erkennen.

Während dieser Arbeit wurde bei der Auswertung der erhaltenen putativen RAS-Sequenzen deutlich, dass einige Sequenzen eher der Sequenz von HST-Enzymen entsprachen. Wie bereits mehrfach erläutert, sind sich RAS- und HST-Enzyme vom Aufbau sehr ähnlich. Im Rahmen dieses Projekts wurden daher mithilfe speziell entwickelter Primer ebenfalls potentielle HST-Sequenzen amplifiziert. Aus der Erfahrung bei der Suche nach RAS-Genen, wurden die verwendeten HST-Primer entsprechend optimiert. Der verwendete 5'-Primer enthält, ähnlich dem RAS-Primer, die konservierte Sequenz HxxxDG. Im Gegensatz dazu enthält der 3'-Primer nicht die Sequenz DFGWG, sondern die leicht abgewandelte Sequenz DFGGX. Die Sequenz wurde aus Betrachtung mehrerer bekannter HST-Sequenzen heraus entwickelt. Nach dem Test mehrerer Sequenzvarianten, stellte sich die Variante DFGGX als beste Sequenz für 3'-Primer heraus.

Die mittels der beschriebenen Primer amplifizierten Teilstücke zeigten auf Aminosäureebene große Ähnlichkeiten zu der Referenzsequenz der HST aus *Glechoma hederacea* (71% bis 100%). Die fünf AS-Sequenzen aus *Alkanna tuberculata*, *Lithospermum officinale*, *Lithospermum purpureocaeruleum*, *Myosotis arvensis* und *Hyssopus officinalis* sind zu 100% identisch (auch auf Ebene der Nukleotidsequenzen) mit der Referenzsequenz aus *Glechoma hederacea*, obwohl vier der fünf Pflanzenarten den Boraginaceae und nicht den Lamiaceae (wie *Glechoma*) angehören.

Die folgende Analyse der erhaltenen AS-Sequenzen der putativen HST-Teilstücke, zeigt deutlich stark konservierte Bereiche. Die konservierten Bereiche am Anfang (Sequenz ab AS 163: HFINTWSDMARGLDITLPPFI) und am Ende der Teilsequenz (Sequenz ab AS 331: YLRSALDFLELQPDLKALVRGAHTFRCPNLGITSWVRLPIH) zeigen dabei große Ähnlichkeiten. Der Großteil der ersten 25 und der letzten 40 Aminosäuren zeigen dabei 100%ige Übereinstimmung. Die miteinander verglichenen Sequenzabschnitte liegen alle außerhalb der verwendeten Primer, daher ist ein versehentlicher Abgleich von Primersequenzen ausgeschlossen.

Zusätzlich sind bei fast jedem Teilstück die Sequenzen VVSIFKLT (K/R) (D/E) Q (L/I) (ab AS 218) (außer bei *Symphytum officinale*) und TKLYIATDGR (ab AS 271) zu finden (die Nummerierung der AS bezieht sich auf die HST-Referenzsequenz aus *Glechoma hederacea*).

Wie bereits in Kapitel 5.2 „Gewinnung von putativen RAS-Teilsequenzen aus verschiedenen Pflanzenarten“ erwähnt, findet man im Weiteren auch hier häufig den Austausch von Aminosäuren mit gleichen chemischen Eigenschaften.

Abbildung 5.4 zeigt die betreffenden Sequenzabschnitte innerhalb eines berechneten 3D-Modells der HST-Sequenz aus *Glechoma hederacea*.

Als erste Gruppe konnten Ma et al. (2004) die Kristallstruktur einer BAHD-Acyltransferase an der Vinorinsynthase aus *Rauwolfia serpentina* nachweisen. Des Weiteren sind HST/HQT Enzyme auch bereits in ihrer Struktur aufgeklärt worden. Kürzlich veröffentlichte Studien führten zu der Aufklärung der exakten dreidimensionalen Struktur (Lallemant et al., 2012a). Sie konnten eindeutig zeigen, dass die Tertiärstruktur der Moleküle ringförmig ausgebildet ist. Das katalytische Zentrum mit dem charakteristischen Motiv HxxxDG liegt innerhalb der Innenseite des Rings.

Wie bereits in Kapitel 5.2 erwähnt, ist es auch hier sehr wahrscheinlich, dass die amplifizierte Teilsequenzen eine identische Struktur innerhalb des Moleküls aufweisen. In Abbildung 5.4 ist ein Argininrest (AS 367, blau) vor dem Eingang des Kanals zum Aktivitätszentrum zu erkennen. Der Argininrest der HST ist stark konserviert und findet sich bei allen Sequenzen, außer bei *Melissa officinalis*. Bei *Melissa officinalis* findet sich dieser Argininrest jedoch auch in der Nähe des DFGWG-Motivs, um zwei Positionen strangaufwärts versetzt (siehe Abbildung 4.9). Wie bereits erwähnt, ist bekannt, dass das Arginin (sowie drei weitere

Aminosäurereste) an dieser Position die Phosphatgruppen des Coenzym-A-Liganden binden (Lallemand et al., 2012b).

Weiter ist innerhalb der putativen Teilsequenzen zu erkennen, dass die folgenden Sequenzabschnitte innerhalb des Kanals bei beiden Enzymen sehr stark konserviert sind:

- RAS ab AS353 QGAHSFR
- HST ab AS351 RGAHTFR

Hier ist vor allem der Austausch des basischen Arginin (HST) zum neutralen Glutamin (RAS) interessant. Diese beiden Aminosäuren sind bei fast allen HST (außer *Symphytum officinale*) und RAS (außer *Melissa officinalis*) innerhalb des Motivs (Q/R)GAH(S/T)FR zu finden. Die im Verlauf dieser Arbeit beobachtete Konservierung dieses Motivs und dessen essentielle Funktion wurden später durch Kristallstudien einer HCT aus *Sorghum bicolor* von Walker et al. (2013) nachgewiesen. Die Bindungstasche wird größtenteils von hydrophoben Aminosäuren gebildet. In der Mitte der Bindungstasche findet sich ein Argininrest an Position 371 (analog zu dem hier gefundenen Argininrest innerhalb der putativen HST Sequenzen an Position 351). Dieser Argininrest hat eine entscheidende Bedeutung bei der enzymatischen Umsetzung von Shikimat. Dies konnten Walker et al. (2013) mit einer HCT-Mutante zeigen, bei der der Argininrest durch ein Alanin ersetzt worden ist (R371A). Die enzymatische Umsetzung von Shikimat ist mit der HCT (R371A)-Mutante nicht mehr nachweisbar.

Die putativen HST Sequenzen in dieser Arbeit enthalten alle einen Argininrest innerhalb des gleichen Sequenzabschnittes kurz vor der DFGWG-Sequenz, der sich im Modell auch mittig innerhalb der Bindungstasche anordnet (Abbildung 5.4). Im Gegensatz dazu enthalten die isolierten putativen RAS-Sequenzen an gleicher Stelle einen konservierten Glutaminrest (Abbildung 5.1).

Die DFGWG-Sequenz an sich ist nicht an der enzymatischen Umsetzung beteiligt (Yu et al., 2009). Es kann somit vermutet werden, dass die genaue Positionierung dieses für die enzymatische Umsetzung wichtigen Argininrestes eine grundlegende Aufgabe der angrenzenden stark konservierte DFGWG-Sequenz ist. Jedoch auch Aminosäurereste hinter dem DFGWG-Motiv sind wichtig für die katalytischen

Eigenschaften von HSTs. So verändert die Mutation von Leu-400-Thr und Phe-402-Tyr die Substratspezifität von Shikimat hin zu Chinat (Lallemant et al., 2012b).

Es zeigt sich, dass bei den gefundenen Sequenzen die Arginin- bzw. Glutaminreste hoch konserviert sind. Walker et al. (2013) beschreiben durch Mutationsstudien die essentiellen Eigenschaften des Argininrestes der HST für die Bindung von Shikimat. Daraus lässt sich ableiten, dass vermutlich auch der neutrale Glutaminrest der RAS die Bindung von 4-Hydroxyphenyllactat beeinflusst bzw. nicht negativ beeinflusst. Die spezifische Aminosäure an dieser Position kann somit als Hinweis zur Charakterisierung für einzelne Hydroxycinnamoyltransferasen dienen.

Befindet sich zum Beispiel kein Argininrest an Position 371, so ist die enzymatische Umsetzung von Shikimat als Akzeptor eher unwahrscheinlich.

In Abbildung 5.2 werden die beiden charakterisierten HST-Sequenzen von *Sorghum bicolor* und *Coffea canephora* verglichen. Das Alignment soll zeigen, ob die von Lallemant et al. (2012b) und Walker et al. (2013) beschriebenen Reaktionsmechanismen und die damit verbundenen Aminosäurereste in beiden Sequenzen relevant sind.



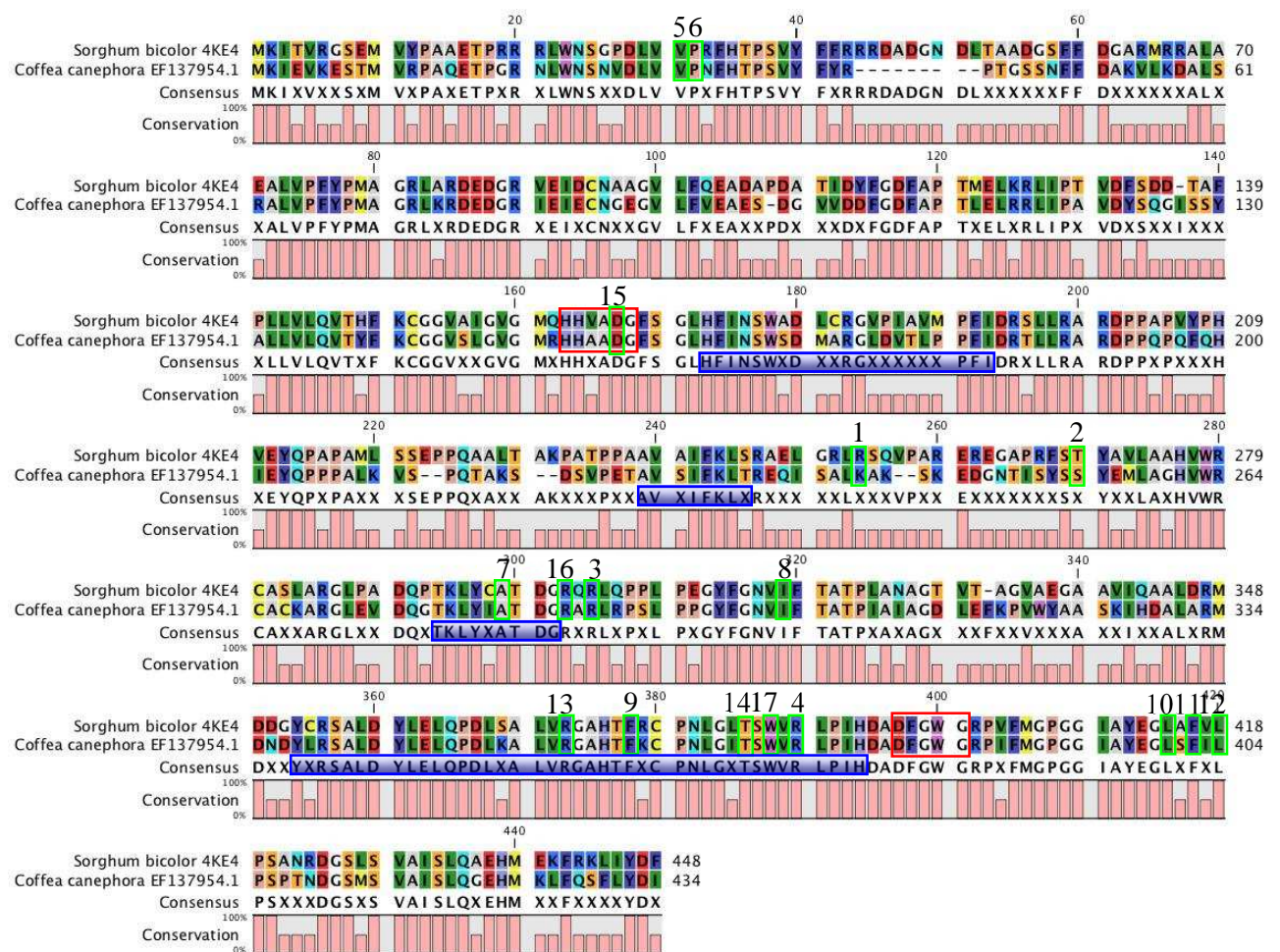


Abbildung 5.2: Alignment der HST-Sequenzen von *Sorghum bicolor* und *Coffea canephora* (Lallemand et al. (2012b) und Walker et al. (2013)). Die charakteristischen Motive HXXXDG und DFGWG sind rot umrahmt. Konservierte Sequenzen, die in den putativen HST Teilsequenzen identifiziert wurden, sind innerhalb der Consensus-Sequenz blau hinterlegt. Die Positionsnummern der Aminosäuren können von denen im Text abweichen, da diese sich durch das Alignment und die eingefügten Gaps verschieben. Die von Lallemand et al. (2012b) und Walker et al. (2013) beschriebenen Reaktionsmechanismen und die damit verbunden Aminosäurereste sind mit einer durchgängigen Zahlenfolge und grünen Kästchen markiert. Die entsprechenden funktionalen Eigenschaften der markierten und nummerierten Aminosäuren werden im Text erläutert.

Die folgenden Sequenzabschnitte wurden von Lallemand et al. (2012b) und Walker et al. (2013) als essentiell für die Durchführung der postulierten Reaktionen bewertet (Die Zahlen hinter den Aminosäureresten markieren die Position in Abbildung 5.2):

1. Die Phosphatgruppen des Coenzym-A-Liganden bindet durch Ionen- oder Wasserstoffbrückenbindungen an folgende Aminosäurereste: Lys-240 (1), Ser-254 (2), Arg-289 (3), und Arg-374 (4) (Lallemand et al., 2012b).

Diese letzten zwei identifizierten Aminosäurereste sind in beiden Sequenzen der HSTs aus *Sorghum bicolor* und *Coffea canephora* identisch. Die Aminosäurereste Lysin 240 und Serin 254 (ausgehend von der Sequenz aus *Coffea canephora*) zeigen hier Unterschiede: Der Aminosäurerest Lysin 240 ist bei der Sequenz von *Sorghum bicolor* durch eine ebenfalls basische Aminosäure, einen Argininrest, ersetzt. Weiter ist im Alignment ersichtlich, dass in der Sequenz von *Sorghum bicolor* ein Serin in unmittelbarer Nähe des Serin 254 lokalisiert ist: Im Alignment um eine Position strangaufwärts versetzt findet sich ein Serinrest in direkter Nachbarschaft (AS269 innerhalb des Alignments).

2. Die Bindungstasche für Shikimat ist überwiegend hydrophob und von Val-31 (5), Pro-32 (6), Ala-298 (7), Ile-318 (8), Phe-376 (9), Leu-414 (10), Phe-416 (11) und Leu-418 (12) umgeben. Im Vergleich mit anderen HCTs sind diese Aminosäurereste nicht stark konserviert, außer dass Sie hydrophobe Eigenschaften aufweisen. In der Mitte der hydrophoben Bindungstasche sind zwei polare Reste Arg-371 (13) und Thr-384 (14) an strategischen Stellen für die Bindung eines Shikimat-Moleküls lokalisiert. Weitere Analysedaten zeigen jedoch eine schwache Interaktion des Shikimats mit den beiden polaren Resten innerhalb der Apo-Form des Enzyms, was auf die weitaus größere Bedeutung der Bindungskräfte bei der Bindung des Coenzym A hinweist (Walker et al., 2013).

Alle zehn oben genannten Aminosäurereste sind in beiden Sequenzen der HSTs aus *Sorghum bicolor* und *Coffea canephora* identisch

3. Die Carboxyl-Seitenkette vom konservierten Rest Asp-166 (15) im HXXXDG Motiv bildet eine Salzbrücke mit dem Arg-302 (16) aus, welches auch stark konserviert ist. Diese Aminosäurereste, Asp-166 und Arg-302, wurden an den jeweiligen

Scheitelpositionen der Domänen I und II lokalisiert und sind sehr wahrscheinlich für die gegenseitige Bindung der Domänen mit verantwortlich (Walker et al., 2013).

Das Arg-302 findet sich im Alignment in beiden Sequenzen der HSTs aus *Sorghum bicolor* und *Coffea canephora*.

4. Beim Katalysemechanismus der HST aus *Sorghum bicolor* dient das His-162 des HXXXDG Motivs als allgemeine Base indem es ein Proton der Shikimisäure während der Katalyse aufnimmt. Das Oxyanion im Tetraederübergangszustand wird durch Wasserstoffbrückenbindungen mit den beiden Seitenketten von Thr-384 (14) und Trp-386 (17) stabilisiert (Walker et al., 2013).

Das Thr-384 und das Trp-386 sind im Alignment der HSTs aus *Sorghum bicolor* und *Coffea canephora* identisch.

Den Aminosäureresten strangabwärts des DFGWG-Motivs ist eine entscheidende Rolle bei der Substratbindung zugeordnet. Auch die Positionierung des DFGWG-Motivs und damit die Koordination der abwärts liegenden Sequenzen innerhalb der Tertiärstruktur können deren Eigenschaften maßgeblich beeinflussen. Innerhalb der entdeckten Teilsequenz einer putativen HST aus *Lavandula angustifolia* konnten die von Lallemand et al. (2012b) und Walker et al. (2013) beschriebenen konservierten Aminosäuren gefunden werden. Hingegen konnten innerhalb der entdeckten Teilsequenz einer putativen RAS aus *Lavandula angustifolia* die konservierten Aminosäuren an manchen Stellen nicht gefunden werden (Abbildung 5.3).

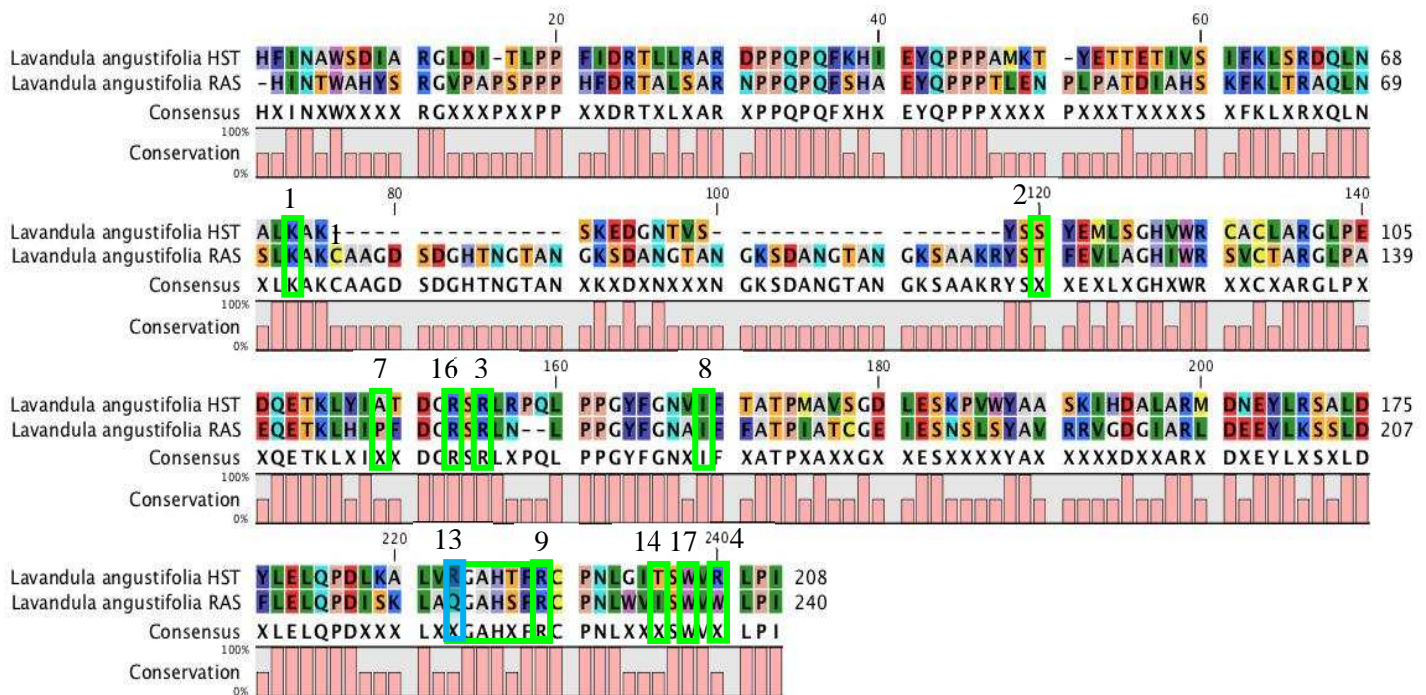


Abbildung 5.3: Vergleich zwischen den putativen RAS und HST Teilsequenzen aus *Lavandula angustifolia*. Es ist zu beachten, dass hier die entdeckten Teilstücke aligniert wurden. Im Weiteren werden jedoch die Positionsnummern der Volllängeensequenzen verwendet, um eine bessere Vergleichbarkeit zu erzielen. Die von Lallemant et al. (2012b) und Walker et al. (2013) beschriebenen Reaktionsmechanismen und die damit verbundenen Aminosäurereste sind mit der zuvor verwendeten durchgängigen Zahlenfolge und grünen Kästchen markiert. Die konservierten Bereiche sind eindeutig auch hier zu erkennen und zu unterscheiden. Die Aminosäuren von RAS und HST sind an folgenden Positionen identisch (1), (16), (3), (8), (9), (17). An folgenden Positionen sind die Aminosäuren unterschiedlich (2), (7), (13), (14), (4). Der von Walker et al. (2013) identifizierte Bereich innerhalb der HST zur Bindung von Shikimat ist gut ersichtlich mit den stark konservierten Resten Arginin Arg-371 (13) und Threonin Thr-384 (14). Hier befindet sich, wie zuvor beschrieben, bei der RAS an gleicher Stelle ein konservierter Glutaminrest (Glu-371). Zusätzlich wird das Oxyanion im Tetraederübergangszustand durch Wasserstoffbrückenbindungen mit den beiden Seitenketten von Thr-384 (14) und Trp-386 (17) stabilisiert. An Position (14) befindet sich bei der RAS jedoch ein Ile-384. Die von Lallemant et al. (2012b) identifizierten Aminosäurereste zur Bindung der Phosphatgruppen des Coenzym-A-Liganden sind identisch (an Positionen (1) und (3)) bzw. ähnlich (an Position (2) findet sich ein Austausch von S/T wie auch zwischen den HSTs aus *Sorghum bicolor* und *Coffea canephora*. Auch hier findet sich ein Serinrest bei der RAS in direkter Nachbarschaft). Einzig an Position (4) ist das Arg-374 durch ein Trp-374 ausgetauscht.

Eine abschließende Klärung dafür, wie sich die Unterschiede bei der RAS auswirken, können dabei nur Mutationsstudien und biochemische Charakterisierungen liefern.

Weiter sind die am stärksten konservierten Sequenzen der putativen HST-Sequenzen, anders als bei der RAS, nicht nur dem Kanal des aktiven Zentrums zuzuordnen (die Sequenz ab AS 271 TKLYIATDGR und die Sequenz ab AS 331 direkt vor dem DFGWG Motiv). Auch die Helix direkt hinter dem aktiven Zentrum HxxxDG ist stark konserviert (Sequenz ab AS 163 hinter HxxxDG Motiv). In Abbildung 5.4 ist dies der rot markierte Bereich unterhalb des grün markierten DFGWG Motivs. Möglicherweise ist die exakte Struktur dieser Helix notwendig, um das aktive Zentrum innerhalb des Moleküls korrekt zu koordinieren. Die Lage der Helix ist im Strukturmodell des Proteins in Abbildung 4.14 zu erkennen.

Die konservierte Sequenz ab AS 218 (VVSIFKLT (K/R) (D/E) Q (L/I)) bildet das erste  $\beta$ -Faltblatt hinter der zuvor beschriebenen ersten Helix direkt hinter dem HxxxDG Motiv. Dieser konservierte Sequenzabschnitt fällt in den Bereich zwischen Pro-206 und Thr-224, der u.a. die beiden Hauptdomänen (Ma et al., 2005) des Moleküls miteinander verbindet und zusammenhält (Lallemant et al., 2012b). In Abbildung 5.4 ist dies der rot markierte Bereich an der Oberseite des Moleküls. Die Lage des  $\beta$ -Faltblatts innerhalb des Moleküls ist auch im Strukturmodell des Proteins in Abbildung 4.14 zu erkennen.



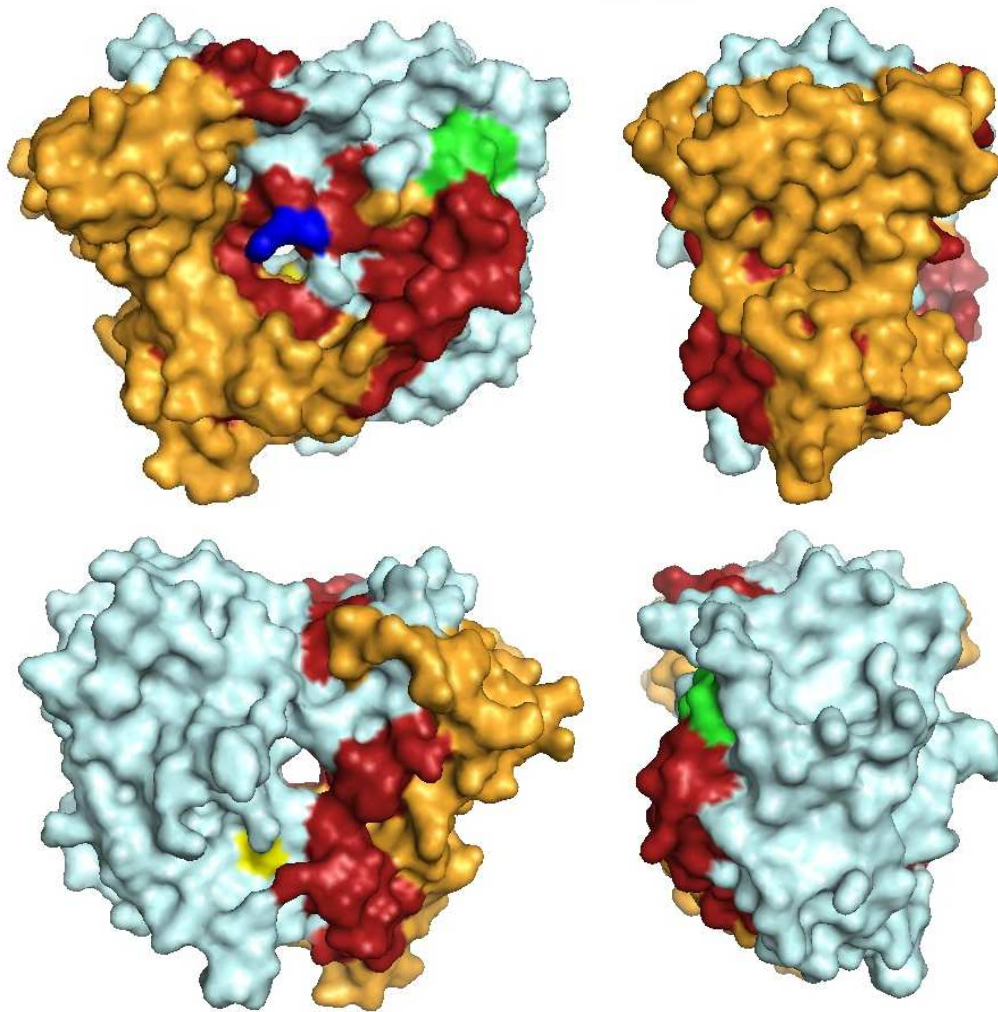


Abb. 5.4: Oberflächendarstellung der HST aus *Glechoma hederacea* mit eingefärbten Sequenzabschnitten. Das Molekül wird in Leserichtung pro Bild um 90° rechtsseitig gedreht. Weiß: nicht amplifizierter Sequenzabschnitt aus der Vollängensequenz von *Glechoma hederacea*, orange: amplifizierter Sequenzabschnitt, rot: stark konservierte Sequenzbereiche, gelb: katalytisches Zentrum mit dem Motiv HxxxDG, grün: charakteristisches Motiv DFGWG, blau: Argininrest, vor dem Kanal zum katalytischen Zentrum angeordnet. Die am stärksten konservierten Sequenzen (rot) sind dem Kanal des aktiven Zentrums und der direkt hinter dem aktiven Zentrum HxxxDG angeordneten Helix zuzuordnen. Die Helix befindet sich unterhalb des grün markierten DFGWG Motivs. Das erste  $\beta$ -Faltblatt hinter dieser Helix ist auch stark konserviert und befindet sich am oberen Rand des dargestellten Moleküls. Die Darstellung wurde mit der frei erhältlichen Education Software Pymol ([pymol.org](http://pymol.org)) erzeugt.



#### **5.4. Gewinnung einer putativen HST-Volllängensequenz aus Beinwell (*Symphytum officinale*)**

Beinwell findet sich in der Flora Europas, Nordamerikas und in Teilen von Asien. Beinwell hat weiter eine lange Tradition als medizinisch wirksame Pflanze. Der Pflanze und ihren Extrakten werden vor allem schmerzstillende und entzündungshemmende Wirkungen nachgesagt und teilweise bestätigt (Kothmann, 2003; Englert et al., 2005). Neben den heilenden Eigenschaften der Pflanze wurden aber auch gesundheitsschädliche Nebenwirkungen festgestellt. Nach Verfütterung von Beinwellwurzeln an Ratten wurde eine starke Lebertoxizität bestätigt. Die Lebertoxizität konnte mit den in der Pflanze enthaltenen Pyrrolizidinalkaloiden (PA) in Zusammenhang gebracht werden (Hirono et al., 1978).

Heute verfügbare medizinische Produkte aus Beinwell enthalten PA-reduzierte oder PA-freie Formulierungen. Bei Anwendung der Produkte wird die maximal tolerierbare Tagesdosis von PA (1 µg/Tag) nicht erreicht. Als Folge gibt es keine Beschränkungen in Deutschland für diese Produkte in Bezug auf die Dauer der Behandlung (Bundesgesundheitsamt, 1992).

Die essentiellen Bestandteile für die medizinisch wirksame Aktivität von Beinwellextrakten und ihre molekularen Wirkungsmechanismen konnten noch nicht vollständig aufgeklärt werden. Allantoin und Rosmarinsäure sind wahrscheinlich von zentraler Bedeutung für die pharmakodynamischen Wirkungen (Andres et al., 1989). Bisher wurden jedoch noch keine klinisch-pharmakokinetischen Untersuchungen am Menschen publiziert.

In dieser Arbeit konnte aus dieser seit langem bekannten Heilpflanze die Volllängensequenz einer noch unbekannten BAHD-Acyltransferase isoliert werden. Die 439 AS lange Sequenz wurde aus der gDNA des Pflanzengewebes von *Symphytum officinale* gewonnen. Mögliche Introns der isolierten DNA-Sequenz, wurden mit dem NetPlantGene Server berechnet. Es konnten keine eindeutigen Introns erkannt werden.

Die Zuordnung der Sequenz zur Familie der Hydroxycinnamoyltransferasen konnte durch die stark konservierten Motive HxxxDG und DFGWG mit Sicherheit bestimmt werden (Abb. 4.12). Durch Sequenzvergleiche wurde zuerst angenommen, dass die gefundene Sequenz in die Gruppe der HST-Enzyme einzuordnen ist. Diese Annahme beruhte auf einem Vergleich verschiedener Aminosäuresequenzen. Die

translatierte Aminosäuresequenz zeigt eine Übereinstimmung von 72% zur HST-Referenzsequenz aus *Glechoma hederacea*, jedoch nur eine Übereinstimmung von 50% zur RAS-Referenzsequenz aus *Coleus blumei*. Werden Aminosäuren mit ähnlichen chemischen Eigenschaften (basisch, hydrophob, aromatisch etc.) in die Berechnung mit einbezogen, ergibt sich eine noch weitaus höhere Übereinstimmung von 82% zur Referenzsequenz aus *Glechoma hederacea*.

Die in dieser Arbeit beschriebene Theorie der möglichen Identifikation von RAS und HST-Sequenzen durch Vergleiche der AS-Positionen innerhalb des Motivs (Q/R)GAH(S/T)FR ist hier nicht möglich. An der beschriebenen Position befindet sich folgendes Motiv:

Ab AS360 KSQNTFK

Wie genau sich der Aminosäurerest Lysin an dieser Position auf die Substratspezifität auswirkt, kann nach momentanem Kenntnisstand nicht beantwortet werden.

Die nahe Verwandtschaft der BAHD-Acyltransferase aus *Symphytum officinale* und der HST-Referenzsequenz aus *Glechoma hederacea* zeigt jedoch auch die mathematische Modellierung der Aminosäuresequenz zur Bestimmung der dreidimensionalen Tertiärstruktur. Für beide Proteine wird eine identische Struktur berechnet, welche eine gute Übereinstimmung der Sekundär- und Tertiärstruktur der Proteine zeigt. Die Proteine enthalten 9 Helices und 6 parallel zueinander angeordnete  $\beta$ -Faltblatt-Strukturen. Weiterhin sind die konservierten Sequenzen HxxxDG und DFGWG identisch angeordnet.

Eine biochemische Charakterisierung ist zur Bestätigung solcher statistischen Aussagen jedoch zwingend erforderlich. Im weiteren Verlauf der Arbeit zeigte sich auch andererseits bei der Berechnung des Phylogramms (Abbildung 5.5) der gefundenen Teilsequenzen, dass die Sequenz aus *Symphytum officinale* eher in die evolutionäre Entwicklung der RAS-Enzyme einzuordnen ist. Die Sequenz hat sich, innerhalb des Phylogramms, sehr früh im evolutionären Kontext innerhalb der RAS-Gruppe entwickelt. Sie ist die erste Sequenz, die sich von den übrigen RAS-Sequenzen an einem eigenen Ast abkoppelt. Sie befindet sich somit auch in direktem verwandtschaftlichem Verhältnis zu der HST-Gruppe und könnte somit

eventuell eine evolutionäre Verbindung zwischen der HST- und der RAS-Gruppe herstellen.

Die Notwendigkeit einer exakten biochemischen Charakterisierung zeigt auch der bereits erwähnte Vergleich der gefundenen BAHD-Acyltransferase aus *Symphytum officinale* mit der RAS-Referenzsequenz aus *Coleus blumei*. Auf Aminosäureebene findet sich eine Übereinstimmung von 50%, sowie eine 70%ige Übereinstimmung von Aminosäuren mit gleicher chemischer Funktionalität. Es ist sogar möglich, dass die isolierte Sequenz HST- als auch RAS-Aktivität zeigt. Eine HQT-Aktivität kann an dieser Stelle eher ausgeschlossen werden, da Beinwell keine Chlorogensäure produziert (Petersen et al., 2009).

Die Expression einer aktiven Form des Enzyms konnte jedoch nicht erzielt werden. Somit konnte die Vermutung, dass es sich bei der gefundenen Sequenz um ein HST-Enzym handelt, nicht endgültig bestätigt werden.

Über die Fehler bei dem Versuch der Expression der BAHD-Acyltransferase kann an dieser Stelle nur spekuliert werden. Folgende für den Erfolg einer Expression abhängigen Variablen wurden vielfach geändert:

- Erntezeitpunkt nach Inkubation
- Inkubationstemperatur
- IPTG-Konzentration
- Induktionszeitpunkt
- Aufschlusspuffer
- Aufschlusszeit
- Expressionszellstamm

Ein weiteres bekanntes Problem ist der Abbau von überexprimierten Proteinen im Cytoplasma (Napoli et al., 1990). Ein rekombinantes Protein kann bei Überexpression sequenzspezifisch abgebaut werden (Que et al., 1997). Dies ist jedoch sehr unwahrscheinlich, da auch bei sehr niedrigen IPTG-Konzentrationen und frühen Erntezeitpunkten kein Protein nachgewiesen werden konnte. Die Menge des vorhandenen Proteins könnte auch unterhalb der Detektionsgrenze der verwendeten Nachweisverfahren gelegen haben.

### **5.5. Erstellung eines phylogenetischen Baums mit den gewonnenen Teilsequenzen der verwendeten Pflanzen**

Für eine phylogenetische Untersuchung kann die Auswahl der Taxa einen großen Einfluss auf das Ergebnis und den weiteren Verlauf der Untersuchungen ausüben. Mit ausreichend Zeit und Mitteln müssten alle Arten eines zu untersuchenden Taxons verwendet werden. Dies war, wie bei den meisten anderen Arbeiten dieser Art, zu umfangreich und nicht umsetzbar. Die getroffene Auswahl beschränkt sich auf die Familien der Boraginaceae und Lamiaceae. Einzelne Vertreter dieser Familien sind bereits für die Synthese von Rosmarinsäure bekannt (Petersen et al., 2009). Hieraus ergibt sich die Frage, ob spezielle Enzyme der Rosmarinsäurebiosynthese in diesen Familienmitgliedern zu finden sind. Könnte man die sie kodierenden Gensequenzen bestimmen, ständen gute molekularbiologische Informationen für kladistische Untersuchungen zur Verfügung. Weiter wurde ein außen stehendes Mitglied der Familie der Blechnaceae untersucht: *Blechnum brasiliense* ist für die Synthese von Rosmarinsäure seit längerem bekannt (Bohm, 1968)

Eine ausreichend hohe Anzahl an unterschiedlichen Taxa ist wichtig, um zu vermeiden, dass symplesiomorphe paraphyletische Einheiten als Monophyla erscheinen (Yoder und Irwin, 1999). Weiter können auch durch scheinbare Analogien paraphyletische Gruppen entstehen, welche Sequenzen mit überdurchschnittlicher Substitutionsrate enthalten. Diese können sehr lange Astabschnitte im Stammbaum (long branch attraction) ausbilden und das Phylogramm verzerren (Felsenstein, 1981).

Die Gesamtzahl der zu untersuchenden Taxa ist jedoch meist unbekannt (Bsp.: Welche Pflanzen enthalten RAS-Enzyme?), so dass nicht bestimmt werden kann, ob die Anzahl der verwendeten Taxa ausreicht, um deren Phylogenie zu rekonstruieren. In dieser Arbeit konnten aus 33 Pflanzen, von insgesamt 49 ausgesuchten Pflanzenarten, entweder eine putative RAS- (insgesamt 12 Pflanzen) oder HST-Teilsequenz (insgesamt 21 Pflanzen) isoliert werden. Aus *Coleus blumei* und *Lavandula angustifolia* konnten beide putativen Sequenzen isoliert werden. Bei der Fülle der verschiedenen Arten innerhalb der Familien der Boraginaceae und Lamiaceae (Freudig und Sauermost, 2004) ist diese Auswahl für eine konkrete phylogenetische Aussage sicherlich unzureichend. Es können aber Tendenzen und

evolutionäre Unterschiede, z.B. zwischen HST und RAS, abgelesen und bewertet werden.

Die Möglichkeit, zu wenige Taxa für die Erstellung eines Phylogramms verwendet zu haben, wird oft als bedeutende Fehlerquelle für phylogenetische Analysen genannt (Lecointre et al., 1993; Yoder und Irwin, 1999). Die Auswirkung einer unzureichenden Anzahl von Taxa auf das Gesamtergebnis wurde mittels Computersimulationen getestet. Auf der Basis realer Datensätze, wurde für eine verminderte Anzahl an Taxa ein wesentlich geringerer Einfluss auf die korrekte Berechnung der Phylogenie gezeigt als z.B. die Anzahl der Nukleotidpositionen oder die Streuung der Substitutionsraten ausüben (Rosenberg und Kumar 2001).

Je mehr Pflanzen bekannt sind, desto genauer kann die evolutionäre Entwicklung bestimmt werden. Einen Ausgangspunkt für die weitere Suche nach geeigneten Taxa zur Erstellung eines phylogenetischen Baums anhand der molekularbiologischen Daten von RAS und HST Sequenzen bietet die hier präsentierte Selektion der Taxa trotz fehlender Isolierung der Gesamtsequenzen. Sollten zukünftig die passenden Gesamtsequenzen der gefundenen Teilsequenzen isoliert werden, kann die hier präsentierte phylogenetische Berechnung zum Vergleich oder als Ausgangsversuch herangezogen werden.

Der momentane wissenschaftliche Erkenntnisstand weist pflanzliche BAHD-Acyltransferasen als monophyletische Gruppe aus. Diese monophyletische Gruppe ist wahrscheinlich aus nicht-pflanzlichen Genen hervorgegangen. Der evolutionäre Ursprung der BAHD-Acyltransferasen wird in Chloramphenicol-Acetyltransferasen (CAT) vermutet (D'Auria, 2006). Diese enthalten ebenfalls das HxxxDG-Motiv, welches grundlegend am Reaktionsmechanismus beteiligt ist. Weiterhin gibt es auch außerhalb des Pflanzenreichs Beispiele von BAHD-Acyltransferasen: Trichothecen 3-O-Acetyltransferasen TRI101 aus *Fusarium sporotrichioides* und *Fusarium graminearum* sind aus einer unbekannten Spezies vermutlich durch horizontalen Gentransfer in das Pilzgenom integriert worden (St. Pierre und De Luca, 2000).

BAHD-Acyltransferasen akzeptieren verschiedenste Substrate, und Enzyme mit neuen Substratakzeptanzen können wahrscheinlich schnell entstehen (Luo et al., 2007). Bei *Arabidopsis thaliana* und *Populus trichocarpa* konnten eine große Anzahl an BAHD-Mitgliedern gefunden werden (Yu et al., 2009). Insgesamt 61 BAHD-Gene liegen über das *Arabidopsis*-Genom verteilt, und 94 Gene weist das *Populus*-Genom auf.

In den Pflanzen wurden paraloge Gene in „tandem repeats“ und auf verschiedenen Chromosomen entdeckt. Dies ist ein Hinweis auf eine verstärkte Genverdopplung. Im *Arabidopsis*-Genom haben mehrmals genomweite Duplikationen und Umorganisationen stattgefunden. Diese haben vermutlich zur Amplifikation der BAHD-Gene und zu weiteren verwandten Enzymen mit verschiedensten Substratspezifitäten geführt (Yu et al., 2009).

Phylogenetische Vergleiche zeigen, dass zu vielen Genen in beiden Genomen keine homologen Gene im jeweils anderen Genom gefunden werden können. Dieses exklusive Vorkommen von BAHD-Genen in einzelnen Pflanzen legt die Vermutung nahe, dass artenspezifische BAHD-Acyltransferasen schnelle evolutionäre Änderungen durchlaufen (Yu et al., 2009).

Innerhalb aller BAHD-Acyltransferasen können eine minimale Sequenzidentität von ca. 10% auf AS-Ebene und die konservierten Motive HxxxD und DFGWG festgestellt werden (Yu et al., 2009). Im Rahmen der Identifizierung von potentiellen BAHD-Acyltransferasen in den Genomen von *Arabidopsis thaliana* und *Populus trichocarpa* entdeckten Yu et al. (2009) jedoch auch, dass das DFGWG-Motiv in *Populus*-Genen regelmäßig wechselt zu DFGFG, DFGWA, DFGWK, NFGWG, DLGFG, DYGWG und NLGWG. Das HxxxD-Motiv fehlt in zwei *Arabidopsis*-Genen und das DFGWG-Motiv variiert hier sogar noch stärker. Von den nachgewiesenen, enzymatisch aktiven BAHD-Acyltransferasen ist das DFGWG-Motiv in insgesamt 14 Fällen zu DFGLG, DFGFG, DFGYG, DFGGG, NFGWG, DFGPG, DFGWD, DFGWN oder DYGWG verändert. Das Motiv DFGGG zeigt dabei eine hohe Ähnlichkeit mit dem in dieser Arbeit verwendeten degenerierten Primer mit DFGGX-Motiv (siehe Kapitel 3.11.5. Konservierte Bereiche der BAHD-Acyltransferasen).

Mit hoher Wahrscheinlichkeit besitzen auch alle BAHD-Acyltransferasen einen einheitlichen Reaktionsmechanismus. Hierfür sprechen außer den konservierten Sequenzmotiven auch die bisher erarbeiteten Kristallstrukturen von BAHD-Acyltransferasen aus drei verschiedenen Klassen (Klasseneinteilung nach D'Auria, 2006):

- Klasse I: Anthocyan Malonyltransferase aus *Dendranthema morifolium*  
(Dm3MAT3, pdb 2e1u, 2e1v) (Unno et al., 2007)
- Klasse III: Vinorinsynthase aus *Rauvolfia serpentina*



(RsVS, pdb 2bgh)

(Ma et al., 2005)

- Klasse VII: Trichothecen Acetyltransferasen aus *Fusarium*-Arten

(FsTRI101/FgTRI101, pdb 3fot, 3fp0) (Garvey et al., 2008)

Alle drei Enzyme haben eine ähnliche Tertiärstruktur. Die Ähnlichkeit der Faltungsmuster ist somit ein weiterer Hinweis auf einen einheitlichen Reaktionsmechanismus.

Das Wissen um Sequenzunterschiede, Tertiärstrukturen und biochemische Eigenschaften der beteiligten Enzyme ist jedoch noch nicht ausreichend geklärt, um genaue Aussagen über die Evolution der Lignin-, Chlorogensäure- und Rosmarinsäurebiosynthesewege zu treffen. Die beteiligten Enzyme katalysieren sehr ähnliche Reaktionen und weisen untereinander eine recht hohe Sequenzähnlichkeit auf. Momentan deuten verfügbare Ergebnisse darauf hin, dass RAS, HST und HQT sehr nah miteinander verwandt sind (Petersen et al., 2009).

Im Weiteren weisen auch die Biosynthesewege große Gemeinsamkeiten auf: Bei der Biosynthese von Caffeoyleshikimat, Chlorogensäure und Rosmarinsäure wird pC-CoA aus dem allgemeinen Phenylpropanstoffwechsel als Hydroxyzimtsäuredonor verwendet. Akzeptoren sind Derivate des Shikimatwegs: Shikimisäure, Chinasäure und 4-Hydroxyphenyllactat. Die Enzyme zur Katalyse der Veresterungen sind alle drei Hydroxyzimtsäuretransferasen aus der Klasse Va der BAHD-Acyltransferasen (Klassifizierung nach D'Auria, 2006). Danach folgt eine *meta*-Hydroxylierung der 4-Cumaroyleinheit, welche von einer Cytochrom P450 Monooxygenase katalysiert wird. Sowohl die Hydroxyzimtsäuretransferasen als auch die *meta*-Hydroxylasen zeigen hohe Sequenzähnlichkeiten und scheinen eine enge evolutionäre Beziehung zu haben, was jedoch noch nicht endgültig bestätigt werden konnte. Für die Substrate ihres jeweiligen Biosynthesewegs sind z.B. Hydroxyzimtsäuretransferasen und CYP98A14 in *Coleus blumei* auch spezifisch, was auf eine evolutionäre Diversifikation der Enzyme hindeutet (Petersen et al., 2009).

Momentan verfügbare Daten zeigen, dass für die Biosynthese von Rosmarinsäure nur wenige neue Enzymaktivitäten entwickelt werden mussten. Diese sind wahrscheinlich aus Genen des Monolignol- oder Chlorogensäuremetabolismus hervorgegangen (Petersen et al., 2009).

Die Ergebnisse dieser Arbeit über die Verwandtschaftsverhältnisse der gefundenen Teilsequenzen können nur eine hypothetische Aussage erbringen. Belastbare Aussagen wären nur mit den entsprechenden Gesamtsequenzen und parallel laufenden biochemischen Charakterisierungen möglich. Mit den RAS- und HST-Teilsequenzen wurde eine Stammbaumanalyse durchgeführt, um einen Überblick zu schaffen und die Sequenzen untereinander zu vergleichen (Abbildung 5.5). Im Folgenden wurde ein phylogenetischer Baum mit den RAS- und HST-Teilsequenzen und weiteren momentan verfügbaren Sequenzen von Hydroxycinnamoyltransferasen aus der pubmed-Proteindatenbank erstellt (Abbildung 5.6).

Der mit den in dieser Arbeit gewonnenen Teilsequenzen konstruierte phylogenetische Baum (Abbildung 5.5) zeigt eindeutig zwei voneinander unabhängige Untergruppen. Eine Gruppe beinhaltet alle gefundenen putativen RAS-Sequenzen und überraschenderweise auch eine putative HST Sequenz aus *Symphytum officinale*. Die betreffende Teilsequenz aus *Symphytum officinale* hat zwar eine starke Ähnlichkeit zur HST-Referenzsequenz aus *Glechoma hederacea* (72%), zeigt aber auch eine ausreichende Übereinstimmung von 50%, sowie eine 70%ige Übereinstimmung von Aminosäuren mit gleicher chemischer Funktionalität mit der RAS-Referenzsequenz aus *Coleus blumei*. Der berechnete phylogenetische Baum aus Abbildung 5.3 ordnet jedoch die Teilsequenz aus *Symphytum officinale* evolutionär den RAS-Sequenzen zu.

Die andere Gruppe enthält alle putativen HST-Sequenzen, die sich ab einem wahrscheinlich frühen evolutionären Zeitpunkt unabhängig von den RAS-Sequenzen entwickelt haben. Auch die putative HST-Sequenz aus *Symphytum officinale* hat sich sehr früh im evolutionären Kontext innerhalb der RAS-Gruppe entwickelt. Sie ist die erste Sequenz, die sich mit einer eigenständigen Entwicklung von den übrigen RAS-Sequenzen abkoppelt. Sie befindet sich somit auch in direktem verwandtschaftlichem Verhältnis zu der HST-Gruppe und könnte eine evolutionäre Verbindung zwischen den beiden Gruppen herstellen. Speziell bei dieser Sequenz aus *Symphytum officinale* wäre eine baldige Feststellung der katalytischen Aktivität wünschenswert.

Der Stammbaum zeigt deutlich, dass durchwegs RAS-Sequenzen und HST-Sequenzen jeweils miteinander gruppieren und nicht die beiden Sequenzen innerhalb ein- und derselben Art näher miteinander verwandt sind.

Generell ist die Sequenzähnlichkeit bei beiden Gruppen in den meisten Fällen sehr hoch (>90 %). Innerhalb der beiden Gruppen, scheinen somit beide Synthesewege

eine alte Fähigkeit darzustellen, die sich in Abhängigkeit von einem gemeinsamen Ursprung entwickelt haben.

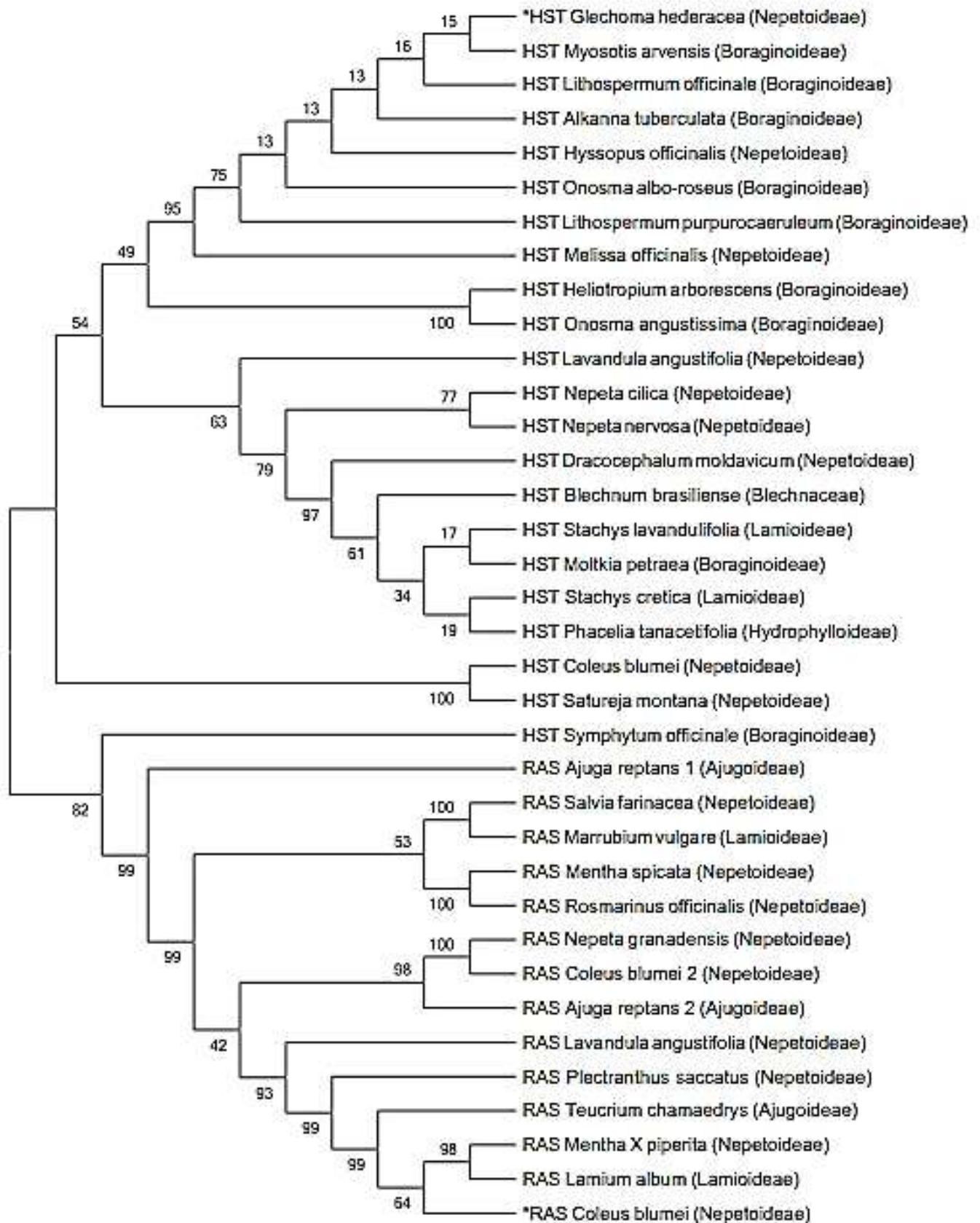


Abb 5.5: Maximum likelihood-Phylogramm (Neighbor-Joining Algorithmus; Pairwise distances berechnet über Jones–Thornton–Taylor Modell für Aminosäuresequenzen) auf der Grundlage des Datensatzes der putativen RAS- und HST-Teilsequenzen. Phylogenetest über Bootstrap Analyse (Anzahl Replikate: 1000). Das Phylogramm wurde mit der Software Mega 5.0 erstellt. Die verwendeten Referenzsequenzen von *Coleus blumei* und *Glechoma hederacea* sind mit einem Stern (\*) gekennzeichnet.

Die Gruppe der in dieser Arbeit isolierten Teilsequenzen wurde um bereits bekannte, veröffentlichte Hydroxycinnamoyltransferase-Sequenzen erweitert (Abbildung 5.6 als A3 Format auf der nächsten Seite angefügt). Es wurden nur eindeutig charakterisierte Proteine aus Samenpflanzen verwendet. Putative Sequenzen wurden ausgeschlossen (siehe Kapitel 3.4.3. BAHD-Acyltransferasen in der pubmed-Datenbank)

Natürlich wurden auch hier ausschließlich die Teilbereiche zwischen den Motiven HxxxDG und DFGWG miteinander verglichen.

Abb 5.6: Maximum likelihood-Phylogramm (Neighbor-Joining Algorithmus; Pairwise distances berechnet über Jones–Thornton–Taylor Modell für Aminosäuresequenzen) auf der Grundlage des Datensatzes der putativen RAS- und HST-Teilsequenzen sowie weiteren Sequenzen in dem Bereich zwischen HxxxDG und DFGWG aus bereits entdeckten Hydroxycinnamoyltransferasen. Phylogenetest über Bootstrap Analyse (Anzahl Replikate: 1000). Das Phylogramm wurde mit der Software Mega 5.0 erstellt. Die während dieser Arbeit isolierten putativen Sequenzen sind folgendermaßen benannt: Ein Stern (\*) wird als Markierung verwendet und danach folgt der Pflanzennamen und anschließend die putative Enzymidentität, z.B.: \**Melissa\_officinalis\_putative\_HST*. Die vollständigen Namen der Hydroxycinnamoyltransferasen sind generell abgekürzt (HQT, HCT, HST, RAS). Bei den Acyltransferasen aus der pubmed-Datenbank wird zuerst die Enzymfamilie und danach der Pflanzennamen in rechteckigen Klammern vermerkt. Folgend ist die alphanumerische Bezeichnung des Proteins eingefügt, z.B.: RAS\_[*Salvia\_miltiorrhiza*].ADA60182. Die entfernt verwandten oder nicht weiter charakterisierten Proteine sind mit vollem Namen ausgeschrieben, z.B.: Anthranilate\_N-hydroxycinnamoyl/ benzoyltransferase\_[*Dianthus caryophyllus*].CAB06427 oder Transferase\_[*Medicago\_truncatula*].ABD28417.

Der phylogenetische Baum wurde in die drei Cluster (CI, CII und CIII) unterteilt. In den jeweiligen Clustern befinden sich die Teilsequenzen, die sich unabhängig voneinander entwickelt haben.

Die putativen HST-Sequenzen aus dieser Arbeit ordnen sich differenziert innerhalb der HQTs, HCTs und HSTs im Cluster CI an. Interessanterweise ordnen sich alle vier HCSQT-Sequenzen innerhalb des CI Clusters an. HCSQTs weisen sowohl HQT- als auch HST-Aktivität auf. Sie ordnen sich im phylogenetischen Baum jedoch innerhalb der putativen HST-Sequenzen und nicht der HQT-Sequenzen an. Somit sollten bei einer biochemischen Charakterisierung der putativen HST-Sequenzen auch Enzyme mit beiden Aktivitätsmerkmalen zu finden sein.

Die putativen RAS-Sequenzen bilden mit den charakterisierten RAS-Sequenzen aus der pubmed-Datenbank ein eigenes Cluster CIII. Innerhalb der RAS-Sequenzen gibt es keine Ausreißer, die sich bei entfernt verwandten Sequenzen anordnen, und sie zeigen im Gesamten eine eigenständige evolutionäre Entwicklung.

Im Gegensatz dazu sind im Cluster CII zwei charakterisierten RAS-Sequenzen aus der pubmed-Datenbank und die putative HST-Sequenzen aus *Symphytum officinale* zu finden. Anhand dieser phylogenetischen Berechnung kann vermutet werden, dass die putative HST aus *Symphytum officinale* sehr früh im evolutionären Kontext eine

eigenständige Entwicklung vollzogen hat. Welchen Einfluss diese Entwicklung hatte, könnte nach Erhalt des aktiven Enzyms eine biochemische Charakterisierung weiter aufklären.

## 6. Zusammenfassung

Pflanzen produzieren eine Fülle von Metaboliten mit geringer molekularer Masse. Nur wenige gehören zum primären Metabolismus, den alle Organismen teilen, der Rest wird dem Sekundärstoffwechsel zugerechnet. Die Biogenese der Sekundärmetabolite ist von der begrenzten Zahl der Primärmetabolite abzuleiten. Weiter sind Sekundärmetabolite zahlreich und in allen höheren Pflanzen zu finden. Dadurch kann das Vorkommen und die Biosynthese von Sekundärstoffen eine phylogenetische Ordnung von Pflanzen ermöglichen.

In dieser Arbeit wurden zwei Enzyme der Biosynthese von Caffeoyleshikimisäure und Rosmarinsäure untersucht. Rosmarinsäure macht den Hauptbestandteil der sogenannten Labiatengerbstoffe aus. Sie besitzt antivirale, antibakterielle und antiinflammatorische Eigenschaften. Caffeoyleshikimisäure hingegen ist eine Vorstufe der Monolignole. Monolignole wiederum sind Grundbausteine des Lignins.

Die Ausgangssubstanzen dieser beiden Sekundärmetabolite sind die Aminosäure L-Phenylalanin und, ebenfalls aus dem Skimatweg, L-Tyrosin bzw. Shikimisäure. An der Biosynthese der Rosmarinsäure sind insgesamt acht Enzyme beteiligt, darunter auch die Hydroxycinnamoyl-CoA:Hydroxyphenyllactat Hydroxycinnamoyltransferase oder Rosmarinsäuresynthase (RAS). An die Stelle der RAS tritt bei der Biosynthese der Monolignole die Hydroxycinnamoyl-CoA:Shikimat Hydroxycinnamoyltransferase (HST). Diese Hydroxycinnamoyltransferasen (HCT) gehören zur Proteinsuperfamilie der BAHD-Acyltransferasen, die im Pflanzenreich ubiquitär verbreitet sind. BAHD-Acyltransferasen sind monomere Enzyme mit einem Molekulargewicht von 48-55 kDa. Die Abkürzung BAHD leitet sich aus den ersten biochemisch charakterisierten Enzymen der Familie ab. HCTs katalysieren sehr ähnliche Reaktionen und weisen untereinander eine recht hohe Sequenzähnlichkeit auf. Fast alle bekannten Mitglieder der BAHD-Familie enthalten das an der Katalyse beteiligte HXXXDG-Motiv im mittleren Bereich des Proteins und das in der Nähe des C-Terminus lokalisierte DFGWG-Motiv.

Im Verlauf der vorliegenden Arbeit konnten putative RAS- und HST-Teilsequenzen aus verschiedenen, nahe verwandten Pflanzenarten gewonnen werden. Die putativen Teilsequenzen haben eine ungefähre Länge von ca. 220 Aminosäuren und befinden sich zwischen den stark konservierten Motiven HXXXDG und DFGWG. Putative RAS-Teilsequenzen wurden ausschließlich aus Pflanzen der Familie der



Lamiaceae isoliert. Putative HST-Teilsequenzen wurden hauptsächlich in Pflanzen der Familien der Lamiaceae und Boraginaceae gefunden. Die Teilsequenzen wurden verwendet, um phylogenetische Stammbäume, u.a. auch zusammen mit bereits bekannten HCTs zu erstellen. Der berechnete phylogenetische Baum wurde in drei Cluster unterteilt. In den jeweiligen Clustern befinden sich die putativen RAS- und HST-Teilsequenzen. Die Ergebnisse deuten darauf hin, dass beide Synthesewege eine alte Fähigkeit darstellen, die sich in Abhängigkeit von einem gemeinsamen Ursprung entwickelt haben. Für eine konkrete phylogenetische Aussage ist die Anzahl der Taxa jedoch sicherlich unzureichend. Es konnten aber Tendenzen und evolutionäre Unterschiede abgelesen und bewertet werden.

Weiter wurden konservierte Sequenzmotive innerhalb der Teilsequenzen identifiziert. Innerhalb der putativen RAS- und HST-Teilsequenzen konnten dabei mehrere unterschiedliche Sequenzmotive identifiziert werden. In Modelldarstellungen ist ersichtlich, dass diese Sequenzmotive hauptsächlich um den Kanal des aktiven Zentrums angeordnet sind. Die konservierten Sequenzmotive der Teilstücke wurden mit bereits bekannten Motiven abgeglichen und Theorien über deren mögliche biochemische Eigenschaft diskutiert.

Außerdem konnte eine bisher unbekannte Volllänge-Sequenz aus Beinwell (*Symphytum officinale*, Boraginaceae) isoliert werden. Hierbei handelt es sich wahrscheinlich um eine Hydroxycinnamoyltransferase, die jedoch nicht exprimiert und biochemisch charakterisiert werden konnte. Die Volllänge-Sequenz wurde in einen evolutionären Kontext gesetzt, *in silico* modelliert und die Daten entsprechend ausgewertet. Daten in Bezug auf Sequenzvergleiche deuteten zuerst auf eine putative HST hin. Im weiteren Verlauf der Arbeit zeigte sich jedoch bei der Berechnung von Phylogrammen unterschiedlicher Sequenzen, dass die Sequenz aus *Symphytum officinale* eher in die evolutionäre Entwicklung der RAS-Enzyme einzuordnen ist. Die Sequenz scheint sich sehr früh im evolutionären Kontext innerhalb der RAS-Gruppe entwickelt zu haben. Sie ist die erste Sequenz, die sich von den übrigen RAS-Sequenzen an einem eigenen Ast abkoppelt. Sie befindet sich somit auch in direktem verwandtschaftlichem Verhältnis zu der HST-Gruppe und könnte eventuell eine evolutionäre Verbindung zwischen der HST- und der RAS-Gruppe herstellen. Hier könnte eine baldige biochemische Charakterisierung sehr viel Klarheit schaffen und damit die getroffenen Annahmen in dieser Arbeit faktenbasiert analysiert werden.

## 7. Ausblick

Der Schwerpunkt der vorliegenden Arbeit lag im Auffinden von RAS-Genen in Pflanzenarten, die nachweislich Rosmarinsäure produzieren. Es konnten putative RAS- und HST- Teilsequenzen in Lamiaceae und Boraginaceae entdeckt werden, die einen Einblick in ihre mögliche, phylogenetische Entwicklung erlauben.

Die hier isolierten Teilsequenzen müssten in Zukunft als Vollängen isoliert werden, um spezifischere Aussagen treffen zu können. Zum Beispiel, um die Frage nach der mono-, para- oder polyphyletischen Entwicklung dieser Hydroxycinnamoyltransferasen weiter aufzuklären.

Weiter wurde eine HST-ähnlichen Gesamtsequenz aus *Symphytum officinale* isoliert, die sich in der phylogenetischen Rekonstruktion der erhaltenen Daten eher den RAS-Sequenzen zuordnet. Die biochemische Charakterisierung jedoch scheiterte.

Sinnvoll erscheint hier die Charakterisierung des funktionsfähigen Enzyms, da erst nach erfolgreichen biochemischen Tests eine exakte Aussage über die Einordnung des Enzyms getroffen werden kann.

Sollte es möglich sein, die gefundenen Teilsequenzen als Vollängen zu isolieren wäre auch hier die biochemische Charakterisierung als erforderliches Mittel zur Bestimmung der Zugehörigkeiten notwendig.

## 8. Literaturverzeichnis

- Agrawal AA, Salminen JP, Fishbein M (2008) Phylogenetic trends in phenolic metabolism of milkweeds (*Asclepias*): Evidence for escalation. The Society for the Study of Evolution Issue Evolution Volume 63 (3): 663–673
- Aharoni A, Keizer LC, Bouwmeester HJ, Sun Z, Alvarez-Huerta M, Verhoeven HA, Blaas J, van Houwelingen AM, De Vos RC, van der Voet H, Jansen RC, Guis M, Mol J, Davis RW, Schena M, van Tunen AJ, O'Connell AP (2000) Identification of the SAAT gene involved in strawberry flavor biogenesis by use of DNA microarrays. Plant Cell (5):647-662
- Andres P, Brenneisen R, Clerc JT (1989) Relating antiphlogistic efficacy of dermatics containing extracts of *Symphytum officinale* to chemical profiles. Planta Med (55): 643-644
- Aquino R, Ciavatta ML, De Simone F, Pizza C (1990) A flavanone glycoside from *Hamelia patens*. Phytochemistry (29): 2358-2360
- Arnold K, Bordoli L, Kopp J, Schwede T (2006). The SWISS-MODEL Workspace: A web-based environment for protein structure homology modelling. Bioinformatics (22): 195-201
- Baltisberger, M (2009) Systematische Botanik. 3. Auflage. vdf Hochschulverlag AG Zürich. ISBN: 9783728131928
- Bentham G (1832-36) Labiatarum Genera et Species. J. Ridgeway & sons, London.
- Bentham G (1876) Labiatae. in: Bentham, G. & Hooker. J.D. (eds.), Genera Plantarum (2) Reeve & Co., Williams & Norgate, London: 1160-1198
- Berg JM, Tymoczko JL (2007) Stryer, Lubert Biochemie 6. Auflage Heidelberg: Spektrum Akademischer Verlag,
- Berger A, Meinhard J, Petersen M (2006) Rosmarinic acid synthase is a new member of the superfamily of BAHD acyltransferases. Planta (224): 1503-1510
- Böttcher F, Adolph RD, Hartmann T (1993) Homospermidine synthase, the first pathway-specific enzyme in pyrrolizidine alkaloid biosynthesis. Phytochemistry (32): 679-689
- Bohm BA (1968) Phenolic compounds in ferns – III. An examination of some ferns for caffeic acid derivatives. Phytochemistry (7): 1825–1830
- Boudet AM (2000) Lignins and lignifications: Selected issues. Plant Physiol Biochem (38): 81-96
- Bourgau F, Gravot A, Milesi S, Gontier E (2001) Production of plant secondary metabolite a historical perspective. Plant Sci (161): 839-851
- Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principles of protein-dye binding. Anal Biochem (72): 248-254

- Bräuchler C (2009) Phylogenetische und taxonomische Untersuchungen an der Subtribus *Menthinae* (Lamiaceae, Nepetoideae, Mentheae) unter besonderer Berücksichtigung des Satureja-Komplexes. Dissertation, LMU München: Fakultät für Biologie. URN: urn:nbn:de:bvb:19-110377. Signatur: 0001/UMC 18324
- Bresinsky A, Körner C, Kadereit JW, Neuhaus G, Sonnewald U (2008) Strasburger Lehrbuch der Botanik. Spektrum Akademischer Verlag. Auflage: 36. ISBN: 3827414555
- Briquet J, (1895-1897) Labiatae. in: Engler, A. & Prantl, K. (eds.), Die natürlichen Pflanzenfamilien (4) Abteilung 3a und 3b. W. Engelmann, Leipzig: 183-380
- Bundesgesundheitsamt (1992) Bekanntmachung über die Zulassung und Registrierung von Arzneimitteln (Abwehr von Arzneimittelrisiken – Stufe II). Bescheid vom 5. Juni 1992, Bundesanzeiger No. 111: 4805
- Burhenne K, Kristensen BK, Rasmussen SK (2003) A new class of *N*-hydroxycinnamoyltransferases. Purification, cloning and expression of a agmatine coumaroyltransferase. *J Biol Chem* (278): 13919-13927
- Cantino PD, Harley RM, Wagstaff SJ (1992) Genera of Labiatae: Status and Classification. *Advances in Labiatae Science Royal Botanic Gardens Kew*: 511-522
- Chase MW, Reveal JL (2009) A phylogenetic classification of the land plants to accompany APG III. *Bot J Linn Soc* (161): 122-127
- Chomczynski P, Sacchi N (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate phenol-chloroform extraction. *Anal Biochem* (162): 156-159
- D'Auria JC (2006) Acyltransferases in plants: a good time to be BAHD. *Curr Opin Plant Biol* (3): 331-340
- De-Eknamkul W, Ellis BE (1987): Tyrosine aminotransferase: the entrypoint enzyme of the tyrosine-derived pathway in rosmarinic acid biosynthesis. *Phytochemistry* (26):1941-1946.
- Dexter R, Qualley A, Kish CM, Je Ma C, Koeduka T, Nagegowda DA, Dudareva N, Pichersky E, Clark D (2007) Characterization of a petunia acetyltransferase involved in the biosynthesis of the floral volatile isoeugenol. *Plant J* (49):265-275
- Doolittle RF (1995) The origins and evolution of eukaryotic proteins. *Philos Trans R Soc Lond B Biol Sci.* (1329):235-40.
- Doolittle WF (1999) Phylogenetic classification and the universal tree. *Science* (5423):2124-2129
- Doyle JJ, Gaut BS (2000) Evolution of genes and taxa: a primer. *Plant Mol Biol* (42):1-23
- Eberle D, Ullmann P, Werck-Reichhart D, Petersen M (2009) cDNA cloning and functional characterisation of CYP98A14 and NADPH:cytochrome P450 reductase from *Coleus blumei* involved in rosmarinic acid biosynthesis. *PlantMol. Biol.* (69): 239-253

- Ellis BE, Towers HG (1970) Biogenesis of rosmarinic acid in *Mentha*. The Biochemical Journal (118): 291-297
- Englberger WU, Hadding E, Etschenberg E, Graf S, Leyck J, Winkelmann M.J (1988) Rosmarinic acid: A new inhibitor of complement C3-convertase with anti-inflammatory activity. International Journal of Immunopharmacology (10): 729-737
- Englert K, Mayer JG, Staiger C (2005) *Symphytum officinale* L.: Comfrey in European pharmacy and medical history. Z Phytother (26): 158-168
- Erdtman G (1945) Pollen morphology and plant taxonomy. IV. Labiatae, Verbenaceae and Avicenniaceae. Svensk Bot. Tidsskr (39): 277-285.
- Farris JS (1970) Methods for Computing Wagner Trees. Syst Zool (19): 83-92
- Felsenstein J (1978): Cases in which parsimony or compatability methods will be positively misleading. Syst Zool (27): 401-410
- Felsenstein J (1981) Evolutionary trees from DNA sequences: a maximum likelihood approach. J Mol Evol (17): 368-376
- Fischbach MA, Walsh CT (2008) The evolution of gene collectives: How natural selection drives chemical innovation. J Proc Natl Acad Sci (105): 4601-4608.
- Fitch WM, Margoliash E (1967) Construction of phylogenetic trees. Science (155): 279 284
- Franke R, Humphreys JM, Hemm MR, Denault JW, Ruegger MO, Cusumano JC, Chapple C (2002) The Arabidopsis REF8 gene encodes the 3-hydroxylase of phenylpropanoid metabolism. The Plant Journal (30): 33-45
- Freudig D, Sauermost R (2004) Lexikon der Biologie. Spektrum Akademischer Verlag ISBN: 3827417368
- Scarpati ML, Oriente G (1958) Isolamento e costituzione dell'acido rosmarinico (dal rosmarinus off.). (28): 2329-2333
- Garvey GS, McCormick SP, Rayment I (2008) Structural and functional characterization of the TRI101 trichothecene 3-O-acetyltransferase from *Fusarium sporotrichioides* and *Fusarium graminearum*: kinetic insights to combating *Fusarium* head blight. J. Biol. Chem. (283): 1660-1669
- Grayer R, Eckert MR, Veitch NC, Kite GC, Marin PD, Kokubun T, Simmonds MSJ, Paton AJ (2003) The chemotaxonomic significance of two bioactive caffeic acid esters, nepetoidins A and B, in the Lamiaceae. Phytochemistry (64): 519-528
- Grayer RJ, Chase MW, Simmonds MSJ (1999) A comparison between chemical and molecular characters for the determination of phylogenetic relationships among plant families: An appreciation of Hegnauer's "Chemotaxonomie der Pflanzen". Biochem Sys Ecol (27): 369-393
- Guex N, Peitsch MC (1997) SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modelling. Electrophoresis (18): 2714-2723

- Harborne JB (1965) Caffeic acid ester distribution in higher plants. *Z Naturforsch* 21b: 604-605
- Harborne JB (1993) Introduction to ecological biochemistry. London Academic Press. ISBN: 0123246865
- Häusler E, Petersen M, Alfermann AW (1991) Hydroxyphenylpyruvate Reductase from Cell Suspension Cultures of *Coleus blumei* Benth. *Z Naturforsch* (46c): 371-376
- Harley RM, Atkins S, Budantsev A, Cantino PD, Conn BJ, Grayer R, Harley MM, De Kok R, Krestovskaja T, Morales R, Paton AJ, Ryding O, Upson T (2004) The Families and Genera of Vascular Plants (7): 167-275
- Hartmann T (1999) Chemical ecology of pyrrolizidine alkaloids. *Planta* (207): 483-495
- Hartmann T (2004) Plant-derived secondary metabolites as defensive chemicals in herbivorous insects: a case study in chemical ecology. *Planta* (219): 1-4
- Hartmann T, Witte L (1995) Chemistry, biology and chemoeology of the pyrrolizidine alkaloids. Pelletier SW, Alkaloids: Chemical and Biological Perspectives (9): 155-233
- Häusler E., Petersen M, Alfermann AW (1992) Rosmarinsäure in *Blechnum*-Spezies. Botanikertagung Berlin. Akademie-Verlag Berlin: 507
- Hebsgaard SM, Korning PG, Tolstrup N, Engelbrecht J, Rouze P, Brunak S (1996) Splice site prediction in *Arabidopsis thaliana* DNA by combining local and global sequence information. *Nucleic Acids Res* (24): 3439-3452
- Hedge IC (1992) A global survey of the biogeography of the Labiatae. *Advances in Labiatae Science* Royal Botanical Gardens Kew London: 7-17
- Hegnauer R (1989) Chemotaxonomie der Pflanzen, vol. 8: Nachträge zu Band 3 und Band 4 (Acanthaceae bis Lythraceae). Birkhäuser, Basel.
- Heldt HW, Piechulla B (2008) Pflanzenbiochemie. Spektrum Akademischer Verlag (4) ISBN: 3827419611
- Hirono I, Mori H, Haga M (1978) Carcinogenic activity of *Symphytum officinale*. *J Natl Cancer Inst* (3): 865-869
- Hoffmann L, Besseau S, Geoffroy P, Ritzenthaler C, Meyer D, Lapierre C, Pollet B, Legrand M (2005) Acyltransferase-catalysed p-coumarate ester formation is a committed step of lignin biosynthesis. *Plant Biosyst* (139): 50-53
- Hoffmann L, Besseau S, Geoffroy P, Ritzenthaler C, Meyer D, Lapierre C, Pollet B, Legrand M (2004) Silencing of hydroxycinnamoyl-coenzyme A shikimate/ quinate hydroxycinnamoyltransferase affects phenylpropanoid biosynthesis. *Plant Cell* (16): 1446-1465



- Hoffmann L, Maury S, Martz F, Geoffroy P, Legrand M (2003) Purification, cloning, and properties of an acyltransferase controlling shikimate and quinate ester intermediates in phenylpropanoid metabolism. *J Biol Chem* (278): 95-103
- Holder M, Lewis PO (2003) Phylogeny estimation: traditional and Bayesian approaches. *Nat Rev Genet* (4): 275-284
- Hughes AL (1994) The evolution of functionally novel proteins after gene duplication. *Proc Biol Sci* (265): 119-124
- Jin L, Nei M (1990) Limitations of the evolutionary parsimony method of phylogenetic analysis. *Mol Biol Evol* (7): 82-102
- Junell, S (1934) Zur Gynäzeummorphologie und Systematik der Verbenaceen und Labiaten, nebst Bemerkungen über ihre Samenentwicklung. *Symbolae Botanicae Upsalienses* (4): 1-219
- Kadereit JW (2004) Flowering Plants · Dicotyledons: Lamiales (except Acanthaceae including Avicenniaceae): Families and Genera of Vascular Plants vol 7 (The Families and Genera of Vascular Plants). Springer ISBN: 3540405933
- Karwatzki B, Petersen M, Alfermann AW (1989) Transient activity of enzymes involved in the biosynthesis of rosmarinic acid in cell suspension cultures of *Coleus blumei*. *Planta Med.* (55): 663-664
- Kempin B (1994) Untersuchungen zur Rosmarinsäuresynthese in Suspensionskulturen von *Coleus blumei*: Tyrosin Aminotransferase. Diplomarbeit, Heinrich-Heine-Universität, Düsseldorf
- Kim KH, Janiak V, Petersen M (2004) Purification, cloning and functiona expression of hydroxyphenylpyruvate reductase involved in rosmarinic acid biosynthesis in cell cultures of *Coleus blumei*. *Plant Mol Biol* (54): 311-323
- Koonin EV (2005) Orthologs, Paralogs and evolutionary Genomics. *Annu Rev Genet* (39): 309-338
- Kothman H (2003) Beinwell - Wirkungsgeschichte und Bedeutungswandel einer Heilpflanze Verlag Dr. Kovac Auflage 1. ISBN: 3830011660
- Kumar S, Tamura K, Nei M (1994) MEGA: Molecular Evolutionary Genetics Analysis software for microcomputers. *Comput Appl Biosci* (10): 189-191
- Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* (227): 680-685
- Lallemand LA, McCarthy JG, McSweeney S, McCarthy AA (2012a) Purification, crystallization and preliminary X-ray diffraction analysis of a hydroxycinnamoyl-CoA shikimate/quinate hydroxycinnamoyltransferase (HCT) from *Coffea canephora* involved in chlorogenic acid biosynthesis. *Acta Cryst Sect F Struct Biol Cryst Commun* (68): 824-828

Lallemant LA, Zubieta C, Lee SG, Wang Y, Acajjaoui S, Timmins J, McSweeney S, Jez JM, McCarthy JG, McCarthy AA (2012b) A structural basis for the biosynthesis of the major chlorogenic acids found in coffee. *Plant Physiol* (160): 249-260

Landmann C, Hücherig S, Fink B, Hoffmann T, Dittlein D, Coiner HA, Schwab W (2011) Substrate promiscuity of a rosmarinic acid synthase from lavender (*Lavandula angustifolia* L.). *Planta* (234): 305-320

Lecointre G, Philippe H, Le H, Le Guyader H (1993) Species sampling has a major impact on phylogenetic inference. *Mol Phylogenet Evol* (2): 205-224

Lindley JC (1829) Notes below *Lophanthus anisatus* description. *Bot. Reg.* (15): 1282

Litvinenko VI, Popova TP, Simonjan AV, Zoz IG, Sokolov VS (1975) Gerbstoffe und Oxyzimtsäureabkömmlinge in Labiaten. *Planta Med* (27): 372-380

Lorence A, Nessler CL (2004) Camptothecin, over four decades of surprising findings. *Phytochemistry* (20): 2735-2749

Luo J, Nishiyama Y, Fuell C, Taguchi G, Elliott K, Hill L, Tanaka Y, Kitayama M, Yamazaki M, Bailey P, Parr A, Michael AJ, Saito K, Martin C (2007) Convergent evolution in the BAHD family of acyl transferases: identification and characterization of anthocyanin acyl transferases from *Arabidopsis thaliana*. *Plant J* (4): 678-695

Ly TN, Shimoyamada M, Yamauchi R (2006) Isolation and characterization of rosmarinic acid oligomers in *Celastrus hindsii* Benth. leaves and their antioxidative activity. *J Agric Food Chem* (54): 3786-3793

Ma X, Koepke J, Panjekar S, Fritzsche G, Stöckigt J (2005) Crystal structure of vinorine synthase, the first representative of the BAHD superfamily. *J Biol Chem* (280): 13576-13583

Maddison WP (1990) A method for testing the correlated evolution of two binary characters. *Evolution* (44): 539-557

Martone PT, Estevez JM, Lu F, Ruel K, Denny MW, Somerville C, Ralph J (2009) Discovery of lignin in seaweed reveals convergent evolution of cell-wall architecture. *Curr Biol* (19): 169-175

Mizukami H, Ellis BE (1991) Rosmarinic acid formation and differential expression of tyrosine aminotransferase isoforms in *Anchusa officinalis* cell suspension cultures. *Plant Cell Rep* (10): 321-324

Molgaard P, Ravn H (1988) Evolutionary aspects of caffeoyl ester distribution in dicotyledons. *Phytochemistry* (27): 2411-2421

Napoli C, Lemieux C, Jorgensen R (1990) Introduction of a chimeric chalcone synthase gene into *Petunia* results in reversible co-suppression of homologous gene in *trans*. *Plant Cell* (2): 279-289

Nei M, Kumar S (2000) Molecular evolution and phylogenetics. Oxford University Press ISBN: 0 195135849

- Niggeweg R, Michael A, Martin C (2004) Engineering plants with increased levels of the antioxidant chlorogenic acid. *Nat Biotechnol* (22): 746-754
- Ober, D, Hartmann, T (1999) Homospermidine synthase, the first pathway-specific enzyme of pyrrolizidine alkaloid biosynthesis, evolved from deoxyhypusine synthase. *Proc Natl Acad Sci* (96), 26: 14777-14782
- Ober D, Harms R, Witte L, Hartmann T (2003) Molecular evolution by change of function. Alkaloid-specific homospermidine synthase retained all properties of deoxyhypusine synthase except binding the eIF5A precursor protein. *J Biol Chem* (278): 12805–12812
- Ohno S (1973) Ancient linkage groups and frozen accidents. *Nature* (244): 259-262
- Pearson K (1900) On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *Philosophical Magazine Series* (302): 157-175
- Pedersen JA (2000) Distribution and taxonomic implications of some phenolics in the family Lamiaceae determined by ESR spectroscopy. *Biochem Syst Ecol* (28): 229-253
- Petersen M, Alfermann AW (1988) Two new enzymes of rosmarinic acid biosynthesis from cell cultures of *Coleus blumei*: hydroxyphenylpyruvate reductase and rosmarinic acid synthase. *Z Naturforsch* (43c): 501-504
- Petersen MS (1991) Characterization of rosmarinic acid synthase from cell cultures of *Coleus blumei*. *Phytochemistry* (9): 2877-2881
- Petersen, M, Häusler E, Karwatzki B, Meinhard J (1993) Proposed biosynthetic pathway for rosmarinic acid in cell cultures of *Coleus blumei*. *Benth Planta* (189): 10-14
- Petersen M (1997) Cytochrome P-450-dependent hydroxylation in the biosynthesis of rosmarinic acid in *Coleus*. *Phytochemistry* (45): 1165-1172
- Petersen M, Simmonds MSJ (2003) Rosmarinic acid. *Phytochemistry* (62): 121-125
- Petersen M, Abdullah Y, Benner J, Eberle D, Gehlen K, Hücherig S, Janiak V, Kim KH, Sander M, Weitzel C, Wolters S (2009) Evolution of rosmarinic acid biosynthesis. *Phytochemistry* (70): 1663-1679
- Pichersky E, Gang DR (2000) Genetics and biochemistry of secondary metabolites in plants: an evolutionary perspective. *Trends Plant Sci* (5): 439-445
- Que Q, Wang HY, English JJ, Jorgensen RA (1997) The frequency and degree of cosuppression by sense chalcone synthase transgenes are dependent on transgene promoter strength and are reduced by premature nonsense codons in the transgene coding sequence. *Plant Cell* (9): 1357-1368
- Rampart M, Beetens JR, Bult H, Herman AG, Parnham MJ, Winkelmann J (1986) Complement-dependent stimulation of prostacyclin biosynthesis: inhibition by rosmarinic acid. *Biochem Pharmacol* (15): 1397-1400

- Razzaque A, Ellis BE (1977) Rosmarinic acid production in *Coleus* cell cultures. *Planta* (137): 287-291.
- Rearick D, Prakash A, McSweeny A, Shepard SS, Fedorova L, Fedorov A (2011) Critical association of ncRNA with introns. *Nucleic Acids Res* (39): 2357-2366
- Reeck GR, de Haën C, Teller DC, Doolittle RF, Fitch WM, Dickerson RE, Chambon P, McLachlan AD, Margoliash E, Jukes TH (1987) Homology in proteins and nucleic acids: a terminology muddle and a way out of it. *Cell* (50): 667
- Reimann A, Nurhayati N, Backenköhler A, Ober D (2004) Repeated evolution of the pyrrolizidine alkaloid-mediated defense system in separate angiosperm lineages. *Plant Cell* (16): 2772–2784
- Rinne, H (2008) Taschenbuch der Statistik. Verlag Harri Deutsch, ISBN: 3817118279
- Rogers SO, Bendich AJ (1985) Extraction of DNA from milligram amounts of fresh, herbarium and mummified tissues. *Plant Mol Biol* (5): 69-76
- Rogozin IB, Sverdlov AV, Babenko VN, Koonin EV (2005) Analysis of evolution of exon-intron structure of eukaryotic genes. *Briefings in Bioinformatics* (6): 118-134
- Rosenberg MS, Kumar S (2001) Incomplete taxon sampling is not a problem for phylogenetic inference. *Proc Natl Acad Sci* (98): 10751-10756
- Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* (4): 406-25
- Sander M (2010) Molekularbiologische und biochemische Untersuchungen von Hydroxycinnamoyltransferasen aus *Coleus blumei* und *Glechoma hederacea*. Dissertation, Philipps-Universität Marburg: Fachbereich Pharmazie. <http://archiv.ub.uni-marburg.de/diss/z2010/0168>
- Sander M, Petersen M (2011) Distinct substrate specificities and unusual substrate flexibilities of two hydroxycinnamoyltransferases, rosmarinic acid synthase and hydroxycinnamoyl-CoA:shikimate hydroxycinnamoyl-transferase, from *Coleus blumei*. *Benth Planta* (6): 1157-1171
- Schaal BA, Hayworth DA (1998) Pylogeographic studies in plants: problems and prospects. *Molecular Ecology* (7): 465-474
- Scheen AC, Bendiksby M, Ryding O, Mathiesen C, Albert V, Lindqvist C. Molecular Phylogenetics, Character Evolution, and Suprageneric Classification of Lamiioideae (Lamiaceae). *Annals of the Missouri Botanical Garden* (97): 191-217
- Schoch GA, Morant M, Abdulrazzak N, Asnaghi C, Goepfert S, Petersen M, Ullmann P, Werck-Reichardt D (2006) The meta-hydroxylation step in the phenylpropanoid pathway. A new level of complexity in the pathway and its regulation. *Environ Chem Lett* (4):127-136

Schoch G, Goepfert S, Morant M, Hehn A, Meyer D, Ullmann P, Werck-Reichhart D (2001) CYP98A3 from *Arabidopsis thaliana* is a 3'-hydroxylase of phenolic esters, a missing link in the phenylpropanoid pathway. *J Biol Chem* (276): 36566-36574

Schwede T, Kopp J, Guex N, Peitsch MC (2003) SWISS-MODEL: an automated protein homology-modeling server. *Nucleic Acids Res* 31: 3381-3385

Shadle G, Chen F, Reddy MSS, Jackson L, Nakashima J, Dixon RA (2007) Down-regulation of hydroxycinnamoyl CoA:shikimate hydroxycinnamoyl transferase in transgenic alfalfa affects lignification, development and forage quality. *Phytochemistry* (68): 1521-1529

Sneath PHA, Sokal R (1973) Numerical taxonomy. The principles and practice of numerical classification. ISBN: 0716706970

Soltis DE, Soltis PS (2003) The Role of Phylogenetics in Comparative Genetics. *Plant Physiol* (132): 1790-1800

Stöckigt J, Zenk MH (1975) Chemical syntheses and properties of hydroxycinnamoyl-coenzyme A derivatives. *Z Naturforsch* (3): 352-358

St. Pierre B, De Luca V (2000) Evolution of acyltransferase genes: origin and diversification of the BAHD superfamily of acyltransferases involved in secondary metabolism. *Rec Adv Phytochem* (34): 285-316

Sullivan ML (2009) A Novel Red Clover Hydroxycinnamoyl Transferase Has Enzymatic Activities Consistent with a Role in Phasic Acid Biosynthesis. *Plant Physiol* (150): 1866-1879

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* (24): 1596-1599

Takeda R, Hasegawa J, Sinozaki M (1990) The first isolation of lignans, megacerotonic acid and anthocerotonic acid, from nonvascular plants, Anthocerotae (hornworts). *Tetrahedron Lett.* (31): 4159-4162

Uduary DW, Zeigler L, Asolkar RN, Singan V, Lapidus A, Fenical W, Jensen PR, Moore BS. (2007) Genome sequencing reveals complex secondary metabolome in the marine actinomycete *Salinispora tropica*. *Proc Natl Acad Sci* 2007 (104): 10376-10381

Unno H, Ichimaida F, Suzuki H, Takahashi S, Tanaka Y, Saito A, Nishino T, Kusunoki M, Nakayama T (2007) Structural and mutational studies of anthocyanin malonyltransferases establish the features of BAHD enzyme catalysis. *J. Biol. Chem.* (282): 15812-15822

Van Tieghem PH (1907) Supplément aux Ochnacées. *Ann Sci Nat Bot* (20): 171-192

Wagner A, Ralph J, Akiyama T, Flint H, Phillips L, Torr K, Nanayakkara B, Te Kiri L (2007) Exploring lignification in conifers by silencing hydroxycinnamoyl-CoA:shikimate hydroxycinnamoyltransferase in *Pinus radiata*. *Proc Natl Acad Sci* (104): 11856-11861

Wagstaff SJ, Olmstead RG, Cantino PD (1995) Parsimony analysis of cpDNA restriction site variation in subfamily Nepetoideae (Labiatae). *Amer J Bot* (82): 886-892

Walker AM, Hayes RP, Youn B, Vermerris W, Sattler SE, Kang C (2013) Elucidation of the structure and reaction mechanism of sorghum hydroxycinnamoyltransferase and its structural relationship to other coenzyme a-dependent transferases and synthases. *Plant Physiol* (162):640-651

Weberling F, Schwantes HO (1975) *Pflanzensystematik. Einführung in die systematische Botanik. Grundzüge des Pflanzensystems*. Verlag Eugen Ulmer Stuttgart. ASIN: B0038ZH2WI

Weitzel C, Petersen M (2011) Cloning and characterisation of rosmarinic acid synthase from *Melissa officinalis* L. *Phytochemistry* (72): 572-578

Wiesmüller B, Henke W, Rothe H (2002) *Phylogenetische Systematik: Eine Einführung*. Berlin/Heidelberg ISBN: 354043643-X

Wink M, Schimmer O (1999) Modes of action of defensive secondary metabolites. In: Wink M, editor. *Functions of plant secondary metabolites and their exploitation in biotechnology*. Blackwell Oxford, *Annual plant reviews* (3): 17–133

Wink M (2003) Evolution of secondary metabolites from an ecological and molecular phylogenetic perspective. *Phytochemistry* (64): 3-19

Wunderlich R (1967) Ein Vorschlag zu einer natürlichen Gliederung der Labiaten aufgrund der Pollenkörner, der Samenentwicklung und des reifen Samens. *Oesterr Bot Z* (114): 383-483

Yoder AD, Irwin AD (1999) Phylogeny of the Lemnaceae: effects of taxon and character sampling on resolution of species relationships within Eulemur. *Cladistics* (15): 351-361

Yu XH, Gou JY, Liu CJ (2009) BAHD superfamily of acyl-CoA dependent acyltransferases in *Populus* and *Arabidopsis*: bioinformatics and gene expression. *Plant Mol Biol* (70): 421-442

## 9. Abkürzungen

|                       |  |
|-----------------------|--|
| <b>% (v/v)</b>        | Volumenprozent   |
| <b>% (m/v)</b>        | Gewichtsprozent  |
| <b>A</b>              | Absorption   |
| <b>Amp</b>            | Ampicillin   |
| <b>APS</b>            | Ammoniumperoxydisulfat   |
| <b>bp</b>             | Basenpaare   |
| <b>BSA</b>            | Rinderserumalbumin (bovine serum albumin)                                  |
| <b>cDNA</b>           | Komplementäre DNA (complementary DNA)                                      |
| <b>CoA</b>            | Coenzym A  |
| <b>Da</b>             | Dalton   |
| <b>DNA</b>            | Desoxyribonucleinsäure (deoxyribonucleic acid)                             |
| <b>dNTP</b>           | desoxy-Nukleosidtriphosphat  |
| <b>DTT</b>            | 1,4-Dithiothreitol   |
| <b><i>E. coli</i></b> | <i>Escherichia coli</i>  |
| <b>EDTA</b>           | Ethylendiamintetraessigsäure   |
| <b>g</b>              | Gramm  |
| <b>h</b>              | Stunden (hours)  |
| <b>HPLC</b>           | Hochdruckflüssigkeitschromatographie (High Pressure Liquid Chromatography) |
| <b>IPTG</b>           | Isopropyl-β-D-thiogalactosid   |
| <b>kat</b>            | Katal  |
| <b>kDa</b>            | Kilodalton   |
| <b>K<sub>m</sub></b>  | Michaelis-Menten-Konstante   |
| <b>l</b>              | Liter  |
| <b>M</b>              | molar (g/mol)  |
| <b>MCS</b>            | multiple Klonierungsstelle (multiple cloning site)                         |
| <b>min</b>            | Minute   |



|                         |   |
|-------------------------|---|
| <b>M<sub>r</sub></b>    | relative Molekülmasse   |
| <b>mRNA</b>             | Boten-RNA (messenger ribonucleic acid)  |
| <b>NMR</b>              | Kernresonanzspektroskopie   |
| <b>OD<sub>600</sub></b> | optische Dichte bei 600 nm  |
| <b>ORF</b>              | offenes Leseraster (open reading frame)   |
| <b>PAGE</b>             | Polyacrylamid-Gelelektrophorese   |
| <b>PCR</b>              | Polymerase Kettenreaktion (polymerase chain reaction)                               |
| <b>PEG</b>              | Polyethylenglykol   |
| <b>RACE</b>             | Verlängerung von cDNA-Enden (rapid amplification of cDNA ends)                      |
| <b>RNA</b>              | Ribonukleinsäure (ribonucleic acid)   |
| <b>RT</b>               | Reverse Transkriptase / Raumtemperatur  |
| <b>SDS</b>              | Natriumdodecylsulfat (sodium dodecylsulfate)  |
| <b>TBE</b>              | Tris Borat EDTA-Puffer  |
| <b>Tris</b>             | Tris-(hydroxymethyl)-aminomethan  |
| <b>UV</b>               | Ultraviolett  |
| <b>Cb</b>               | <i>Coleus blumei</i>  |
| <b>Gh</b>               | <i>Glechoma hederacea</i>   |
| <b>DC</b>               | Dünnschichtchromatographie  |
| <b>DMSO</b>             | Dimethylsulfoxid  |
| <b>FPLC</b>             | Schnelle Protein Flüssigkeits Chromatographie (Fast Protein Liquid Chromatographie) |
| <b>°C</b>               | Grad Celsius  |
| <b>LB</b>               | engl. lysogeny broth  |
| <b>ml</b>               | Milliliter  |
| <b>µl</b>               | Mikroliter  |
| <b>mM</b>               | millimolar  |
| <b>n. d.</b>            | (Reaktionsprodukt) nicht detektierbar   |
| <b>nmol</b>             | nanomolar   |

|                       |  |
|-----------------------|--|
| <b>KP<sub>i</sub></b> | Kaliumphosphatpuffer                                     |
| <b>s</b>              | Sekunde(n)   |
| <b>Tab.</b>           | Tabelle  |
| <b>V</b>              | Volt   |
| <b>U</b>              | Enzymaktivität (Unit) $\mu\text{mol}$ Substratumsatz/min |
| <b>Kana</b>           | Kanamycin  |
| <b>NAD(P)H</b>        | Nicotinamid Adenin Dinukleotid (Phosphat)                |

# Danksagung

Ich möchte nun die Gelegenheit nutzen, um mich bei allen Menschen zu bedanken, die zum Gelingen dieser Arbeit beigetragen haben:

Zunächst möchte ich mich bei Frau Prof. Dr. Maike Petersen für die Möglichkeit zur Erarbeitung dieser Dissertation bedanken. Vielen Dank für die Geduld und Hilfe, die Du mir während der langen Zeit hast stets zukommen lassen.

Herrn Prof. Dr. Michael Keusgen danke ich für die freundliche Übernahme der Zweitkorrektur.

Des Weiteren danke ich recht herzlich Herrn Prof. Dr. Torsten Steinmetzer und Herr Prof. Dr. Christoph Friedrich für die Teilnahme als Prüfer bei meiner Verteidigung.

Ferner danke ich Dr. Nan-Si Brass für die Übernahme der Protokollführung.

Außerdem bedanke ich mich bei allen Kollegen am Institut für Pharmazeutische Biologie für die gute Zusammenarbeit und die angenehme Arbeitsatmosphäre.

Ein besonderer Dank gilt allen Marburgern für die vielen schönen Stunden, als da wären Nan-Si und Sascha mit Mei-Lin und Tian, Frank und Ilka, Evi und Philipp, Diana und Hase, Steffi und Tobi mit Anton und Danny und Jonas.

Mein größter Dank gilt meiner Frau, meinen Eltern und meiner Schwester sowie meiner ganzen Familie, die immer an mich geglaubt haben und ohne deren Unterstützung ich das alles nicht geschafft hätte.

# Erklärung

Ich versichere, dass ich meine Dissertation

## **Phylogenetische Untersuchungen von BAHD-Acyltransferasen in Lamiaceen und Boraginacen**

selbständig und ohne unerlaubte Hilfe angefertigt und mich dabei keiner anderen als der von mir ausdrücklich bezeichneten Quellen bedient habe.

Die Dissertation wurde in der jetzigen oder ähnlichen Form noch bei keiner anderen Hochschule eingereicht und hat noch keinen sonstigen Prüfungszwecken gedient.

Herbornseelbach, den 12. November 2014

(Johannes Benner)

# Lebenslauf

Johannes Michael Benner  
geb. am 03. Dezember 1979  
in Herborn (Dillkreis)  
Staatsangehörigkeit: deutsch

## Schulbildung:

|           |   |
|-----------|---|
| 1999      | Abschluss: Allgemeine Hochschulreife    |
| 1990-1999 | Johanneum-Gymnasium der Stadt Herborn   |
| 1986-1990 | Dernbach-Grundschule in Herbornseelbach |

## Zivildienst:

|                         |   |
|-------------------------|---|
| 08 / 1999 bis 09 / 2000 | Mobile soziale Hilfsdienste bei der Diakonie in Wetzlar |
|-------------------------|---|

## Studium:

|                         |   |
|-------------------------|---|
| 01 / 2008 bis 12 / 2010 | Wissenschaftlicher Mitarbeiter am Institut für Pharmazeutische Biologie und Biotechnologie der Philipps-Universität Marburg |
| 12 / 2007               | 3. Staatsexamen und Approbation als Apotheker   |
| 05 / 2007 bis 10 / 2007 | Pharmaziepraktikum in der Apotheke am Obermarkt in Frankenberg (Eder)   |
| 11 / 2006 bis 04 / 2007 | Pharmaziepraktikum am Institut für Pharmazeutische Chemie der Philipps-Universität Marburg                                  |
| 10 / 2002 bis 09 / 2006 | Pharmaziestudium an der Philipps-Universität Marburg  |
| 10 / 2000 bis 09 / 2002 | Studium der Mathematik (Diplom) ohne Abschluss an der Justus-Liebig-Universität Gießen                                      |

## Berufspraxis:

|                         |   |
|-------------------------|---|
| 12 / 2007 bis 12 / 2010 | Nebentätigkeit in der Apotheke am Obermarkt in Frankenberg (Eder) |
|-------------------------|---|

## Selbstständigkeit:

|                |  |
|----------------|--|
| Seit 05 / 2011 | Inhaber der Eder-, Kloster- und Löwen-Apotheke in Frankenberg (Eder) |
|----------------|--|

Frankenberg (Eder) den 12. November 2014



